

Bachelorarbeit im Studiengang
Bibliotheks- und Informationsmanagement

**Transliterationsfehler bei der Formalerschließung
russischsprachiger Monografien
in der Verbunddatenbank K10plus**

vorgelegt von
Anna Seeger
an der Hochschule der Medien Stuttgart

zur Erlangung des akademischen Grades eines Bachelor of Arts

Erstprüferin: Prof. Heidrun Wiesenmüller
Zweitprüferin: Silke Horny

Ehrenwörtliche Erklärung

Hiermit versichere ich, Anna Seeger, ehrenwörtlich, dass ich die vorliegende Bachelorarbeit mit dem Titel: „Transliterationsfehler bei der Formalerschließung russischsprachiger Monografien in der Verbunddatenbank K10plus“ selbstständig und ohne fremde Hilfe verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Stellen der Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen wurden, sind in jedem Fall unter Angabe der Quelle kenntlich gemacht. Die Arbeit ist noch nicht veröffentlicht oder in anderer Form als Prüfungsleistung vorgelegt worden.

Ich habe die Bedeutung der ehrenwörtlichen Versicherung und die prüfungsrechtlichen Folgen (§26 Abs. 2 Bachelor-SPO (6 Semester), § 24 Abs. 2 Bachelor-SPO (7 Semester), § 23 Abs. 2 Master-SPO (3 Semester) bzw. § 19 Abs. 2 Master-SPO (4 Semester und berufsbegleitend) der HdM) einer unrichtigen oder unvollständigen ehrenwörtlichen Versicherung zur Kenntnis genommen.

Ort, Datum

Unterschrift

Abstract

In dieser Bachelor-Arbeit wurden Transliterationsfehler bei der Formalerschließung russischsprachiger Monografien in der Verbunddatenbank K10plus untersucht. Der Fokus der Untersuchung lag dabei auf den Feldern Haupttitel und Titeltzusatz, bei denen üblicherweise großer Transliterationsbedarf besteht und der Aufwand zu derer Überprüfung überschaubarer war. Die Daten wurden in der Verbunddatenbank mittels speziell formulierter Suchanfrage für die Jahrgänge 2011-2015 abgerufen und überprüft. Festgestellte Transliterationsfehler wurden nach Fehlerarten gruppiert und in Bezug auf ihre Häufigkeit, Vorkommen bestimmter Muster und sonstige Auffälligkeiten analysiert. Gesondert wurde ihre mögliche Auswirkung auf das Retrieval beschrieben. Anschließend wurden mögliche Hilfestellungen für die Praxis formuliert. Sämtliche geprüften Datensätze und festgestellten Fehler wurden in einem Analyseprotokoll festgehalten, welches als Anhang zur Bachelor-Arbeit in elektronischer Form aufgeführt wird.

Schlagwörter: Katalogisierung; Verbundkatalog; Transliteration; Kyrilliza

This thesis deals with transliteration errors that occur while cataloguing Russian-language monographs in the cooperative cataloguing database K10plus. The analysis deals with the cataloguing fields main title and other title information, which usually require a lot of transliteration and are easier to check within the scope of a Bachelor thesis. The data is retrieved and checked using a specially formulated search query in the K10plus database for the years 2011-2015. Detected transliteration errors are grouped according to error types and analysed concerning their frequency, occurrence of certain patterns and other abnormalities. Their possible impact on retrieval is described separately. Subsequently, possible practical help is formulated. All checked data records and identified errors are recorded in an analysis protocol, which is listed as an appendix to the thesis in electronic form.

Keywords: cataloguing; cooperative cataloguing database, transliteration, Cyrillic script

Inhaltsverzeichnis

Abbildungsverzeichnis.....	5
Tabellenverzeichnis.....	6
1 Einleitung.....	7
2 Hintergründe und Stand der Forschung.....	10
2.1 Transliteration	10
2.1.1 Begriffsklärung Transliteration und Transkription.....	10
2.1.2 Transliterationssysteme für die russische Schrift.....	11
2.1.3 Regelungen für bibliothekarische Transliteration in Deutschland.....	15
2.1.4 Regelungen für die Erfassung nicht-lateinischer Schriften in K10plus	19
2.1.5 Kritik an Transliteration in bibliothekarischen Datenbanken.....	19
2.2 OPAC-Qualität und Tippfehler	21
2.2.1 Fehlerfreier OPAC als Qualitätskriterium.....	21
2.2.2 Tippfehler in bibliothekarischen Datenbanken	22
3 Analyse der Tippfehler in K10plus	26
3.1 Beschreibung des Forschungsvorhabens.....	26
3.1.1 Forschungsfragen	26
3.1.2 Methodik.....	26
3.1.3 Einschränkungen.....	29
3.2 Ergebnisse der Analyse.....	30
3.2.1 Gesamtzahl der ermittelten und geprüften Datensätze	30
3.2.2 Gesamtzahl der fehlerhaften Datensätze und der festgestellten Fehler.....	30
3.2.3 Aufschlüsselung der Datensätze nach besitzender Bibliothek	31
3.2.4 Aufschlüsselung der festgestellten Fehler nach Fehlerart.....	34
3.3 Beantwortung der Forschungsfragen und Zusammenfassung der Analyseergebnisse	66
3.4 Mögliche Hilfestellungen.....	68
4 Fazit und Ausblick.....	71
Literaturverzeichnis	73
Anhang A. Analyseprotokoll	78

Abbildungsverzeichnis

Abbildung 1. Anteile der einzelnen Fehlerarten an der Gesamtfehlerzahl.....	35
Abbildung 2. Buchcover zum Datensatz PPN 1645265862	42
Abbildung 3. Titelblatt zum Datensatz PPN 795228880.....	44
Abbildung 4. Buchcover zum Datensatz PPN 161883679X.....	48
Abbildung 5. Titelblatt zum Datensatz PPN 161883679X	49
Abbildung 6. Rückseite des Titelblattes zum Datensatz PPN 161883679X.....	49
Abbildung 7. Beispiel einer Titelseite nach GOST R 7.0.4.-2006	52
Abbildung 8. Buchcover und Titelblatt zum Datensatz PPN 783671687	55

Tabellenverzeichnis

Tabelle 1. Vergleich ausgewählter Transliterationsstandards	14
Tabelle 2. Vergleich deutscher bibliothekarischer Transliterationsschemas	17
Tabelle 3. Anzahl der ermittelten und geprüften Datensätze nach Jahrgang	30
Tabelle 4. Verteilung der fehlerhaften Datensätze auf jeweilige Jahrgänge	31
Tabelle 5. Aufschlüsselung der Datensätze nach besitzenden Bibliotheken	32
Tabelle 6. Aufschlüsselung der festgestellten Fehler nach Fehlerart	34
Tabelle 7. Vertauschung in der Stichprobe	36
Tabelle 8. Beispiele von Auslassungen	37
Tabelle 9. Ausgelassene Buchstaben und Häufigkeit ihrer Auslassung	38
Tabelle 10. Auslassung des Buchstabens „б“ nach erfassender Bibliothek	39
Tabelle 11. Beispiele der Weglassung diakritischer Zeichen	40
Tabelle 12. Aufschlüsselung der Weglassung diakritischer Zeichen	41
Tabelle 13. Beispiele für das Einfügen	43
Tabelle 14. Beispiele für Ersetzungen	45
Tabelle 15. Aufschlüsselung der Ersetzungen	46
Tabelle 16. Aufschlüsselung der Groß- und Kleinschreibfehler	57
Tabelle 17. Verwendung falscher diakritischer Zeichen	60
Tabelle 18. Sonstige Auffälligkeiten, Teil 1	61
Tabelle 19. Sonstige Auffälligkeiten, Teil 2	63
Tabelle 20. Sonstige Auffälligkeiten, Teil 3	65

1 Einleitung

Der Begriff Transliteration stammt aus dem Bereich der Sprachwissenschaft. Wie die Begriffsklärung später zeigen wird, scheinen der Begriff selbst und die damit verbundene Vorgehensweise relativ klar zu sein – man nehme ein Wort aus einer Sprache mit einer bestimmten Schrift und schreibe es auf Grundlage einer fest definierten Transliterationstabelle mit Buchstaben oder Buchstabenkombinationen einer anderen Sprache um.

Wie in vielen anderen Bereichen sieht die Praxis leider etwas anders aus. Einer der Gründe dafür liegt darin, dass weltweit mehrere Transliterationsansätze und somit mehrere Transliterationsstandards existieren und angewendet werden. Somit hat die Transliterationsproblematik unmittelbare Auswirkungen weit außerhalb der sprachwissenschaftlichen Kreise. So verhandelt der Europäische Gerichtshof für Menschenrechte regelmäßig Streitigkeiten, die Transliteration von Personennamen zum Gegenstand haben.¹ Auch der Verfasserin dieser Bachelor-Arbeit, die als russische Staatsangehörige in Deutschland lebt, blieben ähnliche Transliterationsprobleme nicht erspart.

An dieser Stelle möge man meinen, dass diese Problematik im Bereich Bibliothekswesen weniger relevant ist, denn bibliothekarische Standards im deutschsprachigen Raum (RDA und dazugehörige D-A-CH-Empfehlungen) definieren klare Transliterationsregeln. So ist im RDA-Toolkit die Tabelle für die Umschrift kyrillischer Alphabete slawischer Sprachen enthalten, die für eine einheitliche Transliteration im bibliothekarischen Bereich sorgt. Jedoch kann das Problem der unterschiedlichen Transliterationsstandards spätestens bei der Übernahme von Fremddaten wieder ans Licht kommen. Je nach Anbieter der Fremddaten kann es sein, dass andere Transliterationsregeln verwendet werden. Eine gewisse Herausforderung stellen auch Übersetzungen der russischsprachigen Werke in andere Sprachen dar. Die Titel solcher Werke werden üblicherweise übersetzt, liegen somit in der Sprache der Übersetzung vor (z. B. Deutsch oder Englisch) und brauchen meistens gar nicht transliteriert zu werden. Allerdings wurden in solchen Fällen im Laufe der Übersetzung die Personen- und Körperschaftsnamen bereits transliteriert, und zwar nach den Vorstellungen der Übersetzer/innen oder nach Vorgaben des Verlages und womöglich nach ganz anderen Transliterationsstandards. Als Folge könnte es dazu kommen, dass anstatt der Verlinkung eines bestehenden Normdatensatzes eine Dublette mit einer anderen Transliterationsvariante angelegt wird. Wenn tatsächlich kein entsprechender Normdatensatz vorhanden ist, entsteht ein neuer Normdatensatz mit einem russischen Namen, der allerdings nach anderen Standards transliteriert ist.

¹ Vgl. European Court of Human Rights (2007): Case of Bulgakov v. Ukraine, Application no. 59894/00; European Court of Human Rights (2004): Case of Mentzen v. Latvia, Application no. 71074/01; Cour européenne des Droits de l'Homme (2004): L'affaire Kuharec alias Kuhareca c. Lettonie, requête no 71557/01; Cour européenne des Droits de l'Homme (2001): L'affaire Siskina and Siskins v. Latvia, requête no 59727/00 – dieser Fall erregte die meiste Aufmerksamkeit in den russischen Medien und Übersetzerkreisen, vgl. dazu: Bologov, Petr (2013): Vse, ja teper' Sis'kin!; Ermolovič, D.I. (2012) Šiškin protiv Siskina.

Falsche bzw. nicht einheitliche Transliteration von Namen sowie das öfter vorkommende Wegfallen der bei Transliteration verwendeten diakritischen Zeichen während der weiteren Datenverarbeitung können außerhalb der Formalerschließung für wissenschaftliche Bibliotheken relevant sein, die in den Bereichen wissenschaftliches Publizieren, Publikationssupport und Bibliometrie tätig sind. Transliterationsprobleme könnten zur Entstehung mehrerer Namensvarianten und somit zur falschen Zuordnung der Zitate und Verzerrung der bibliometrischen Indikatoren führen, die ihrerseits zur Bewertung einzelner Wissenschaftler*innen und ganzer Forschungseinrichtungen verwendet werden. In bibliothekarischen Katalogen wird diese Problematik durch die Verwendung der Normdaten gelöst, führende wissenschaftliche Zitationsdatenbanken wie Web of Science und Scopus haben allerdings mit vielen Mutationen sogar bei vermeintlich einfacher zu verarbeitenden europäischen Namen zu kämpfen.² Probleme mit der Schreibweise von Namen mit diakritischen Zeichen sind somit geradezu vorprogrammiert. Diese Angelegenheit gehört allerdings hauptsächlich in den Kompetenzbereich der Datenbankanbieter und nicht der wissenschaftlichen Bibliotheken.

Da die Bestände an russischsprachigen Werken in deutschen Bibliotheken relativ beschränkt sind, kommen die beschriebenen, Bibliotheken betreffende Problemfälle wahrscheinlich nicht so oft vor und stellen dementsprechend keine so große Herausforderung für die deutschen OPACs dar. Im Rahmen ihres Praktikums in der Bibliothek des Staatlichen Museums für Naturkunde Stuttgart ist der Verfasserin dieser Bachelor-Arbeit ein anderes Problem aufgefallen – bei der Formalerschließung russischsprachiger Medien kommen (Tipp-)Fehler gefühlt etwas öfter vor, als man es sich bei deutsch- und englischsprachigen Medien gewohnt ist. Da diese Feststellung jedoch auf einer kleineren Zahl der Datensätze basierte, lässt sich nicht beurteilen, ob diese Tatsache der Wirklichkeit entspricht oder an einer etwas unglücklicheren Zusammenstellung der einzelnen Datensätze basiert. In diesem Moment entstand die Idee, derartigen Tippfehlern im Rahmen einer Bachelor-Arbeit detaillierter nachzugehen und herauszufinden, wie oft diese transliterationsbedingten Tippfehler tatsächlich vorkommen, ob sich dabei Muster identifizieren lassen und an welchen Stellen die meisten Probleme entstehen.

Die vorliegende Bachelor-Arbeit gliedert sich inhaltlich in zwei Teile. Zuerst werden im Kapitel 2 die Hintergründe zur Transliteration beschrieben. Der Begriff Transliteration wird definiert, es wird auf verschiedene Transliterationssysteme auf internationaler und nationaler Ebene eingegangen. Danach widmet sich Kapitel 2 Tippfehlern in bibliografischen Datenbanken. Hier wird die Rolle der Tippfehler für die Qualitätsbeurteilung einer bibliografischen Datenbank geprüft sowie der Stand der Forschung zu Tippfehlern in diesem Bereich beschrieben.

Kapitel 3 stellt den praktischen Teil der Bachelor-Arbeit dar. Nach der Beschreibung des Forschungsvorhabens und Formulierung der Forschungsfragen wird auf die Methodik

² Vgl. Tüür-Fröhlich, Terje (2018): Eine „autoritative“ Datenbank auf dem Prüfstand: Der Social Sciences Citation Index (SSCI) und seine Datenqualität, S. 268; sowie Franceschini, Fiorenzo; Maisano, Domenico; Mastrogiovanni, Luca (2016): Empirical analysis and classification of database errors in Scopus and Web of Science, Seite 942.

und ihre möglichen Einschränkungen eingegangen. Danach werden die Ergebnisse der durchgeführten Analyse vorgestellt und detailliert beschrieben. Anschließend erfolgt ihre Zusammenfassung. Antworten auf die im Anfang des Kapitels 3 gestellten Forschungsfragen werden formuliert.

Im letzten Kapitel „Fazit und Ausblick“ werden die wichtigsten Erkenntnisse der Arbeit nochmals zusammengefasst sowie ein Ausblick für die weitere Forschung dargestellt.

2 Hintergründe und Stand der Forschung

2.1 Transliteration

2.1.1 Begriffsklärung Transliteration und Transkription

Der Begriff Transliteration stammt aus der Sprachwissenschaft. Ein verwandter Vorgang ist dabei die Transkription. In der Praxis kommt es jedoch zu einer Verwechslung der beiden Begriffe³, daher sollen als Erstes die Unterschiede zwischen den beiden Vorgängen definiert werden. Dafür werden jeweils zwei Definitionen verwendet, da diese sich gegenseitig ergänzen und zusammen betrachtet die Unterschiede zwischen Transkription und Transliteration besser hervorheben lassen.

Transkription bedeutet laut der Definition des Lexikons der Sprachwissenschaft

„Vorgang und Ergebnis der Wiedergabe eines Textes beliebiger Verschriftung [...] in Form eines alphabetischen Textes. Bei T[ranskription] besteht zwischen transkribiertem und transkribierendem Text in der Regel keine Eins-zu-eins-Beziehung. So entspricht sch in der Duden-T[ransliteration] des Russ[ischen] entweder <Ш> oder <Ж>.“⁴

Die zweite Definition laut dem Metzler Lexikon Sprache bringt einen weiteren wichtigen Aspekt hinein, der den Transkriptionsvorgang ausmacht, und zwar:

„[die] Wiedergabe der Aussprache von lexikal[ischen] Einheiten oder Texten bzw. tatsächl[ich] vorliegender lautsprachtl[icher] Äußerungen [...].“⁵

Der Begriff Transliteration bezeichnet laut dem Lexikon der Sprachwissenschaft

„Vorgang und Ergebnis der uneindeutigen Übertragung eines in alphabetischer oder syllabischer Schrift geschriebenen Textes in einen alphabetischen Text, der die Gestalt des transliterierten Textes eindeutig zu rekonstruieren gestattet.“⁶

Die zweite Definition des Metzler Lexikons Sprache lautet wie folgt:

„Umsetzen eines Textes aus einer Schriftart in eine andere, z. B. von japan[ischen] (silbischen) Schriftzeichen in eine lat[einische] (alphabet[ische]) Buchstabenfolge oder zwischen kyrill[isch] und lat[einisch] geschriebenen Texten. Für die T[ransliteration] bestehen verschiedene internat[ionale] und nat[ionale] Standardisierungen.“⁷

³ Vgl. Glück, Helmut, Hrsg.; Rödel, Michael, Hrsg. (2016): Metzler Lexikon Sprache, S. 723.

⁴ Bußmann, Hadumod, Hrsg. (2008): Lexikon der Sprachwissenschaft, S. 752.

⁵ Glück, Helmut, Hrsg.; Rödel, Michael, Hrsg. (2016): Metzler Lexikon Sprache, S. 722.

⁶ Bußmann, Hadumod, Hrsg. (2008): Lexikon der Sprachwissenschaft, S. 752.

⁷ Glück, Helmut, Hrsg.; Rödel, Michael, Hrsg. (2016): Metzler Lexikon Sprache, S. 723.

Aus den aufgeführten Definitionen lassen sich die wichtigsten Unterschiede zwischen Transkription und Transliteration ableiten. Der erste Vorgang orientiert sich an der Aussprache, der zweite hingegen an der genauen Folge der Schriftzeichen oder Buchstaben. Dies hat entscheidende Folgen für das Ergebnis der beiden Vorgänge. Bei einem transliterierten Text ist es möglich, den Ausgangstext eindeutig zu rekonstruieren. Bei einem transkribierten Text ist eine solche Rekonstruktion nicht durchführbar.

Der Unterschied zwischen Transliteration und Transkription lässt sich durch das folgende Beispiel veranschaulichen. Das russische Wort „защита“ („beschützende“) wird ins Englische wie „zashchishchayushchaya“ und ins Deutsche wie „sachtschischtschajuschschaja“ transkribiert, wodurch der Leser, der kein Russisch spricht, basierend allein auf der allgemeinen Orthografie der eigenen Sprache die (zumindest ungefähre) Aussprache des russischen Wortes nachahmen kann. Die Transliteration nach DIN 1460 wäre in diesem Fall „zaščičajuščaja“.⁸

Das deutsche Wort „Umschrift“ steht laut Duden für die phonetische Umschrift oder Transkription⁹, wird in der Praxis aber ebenfalls für Transliteration verwendet.¹⁰ Da es in der Praxis jedoch kaum „reine“ Transliterations- oder Transkriptionssysteme gibt und viele Systeme bestimmte Kompromisse eingehen¹¹, lehnt sich die vorliegende Arbeit, insbesondere im Punkt 2.1.2, an die Definition an, wie sie von Wellisch verwendet wird. Wellisch versteht den Begriff Umschrift entweder umfassend als Transliteration und Transkription oder als einen Vorgang, der beide Verfahren gleichzeitig kombiniert, denn in der Praxis ist es sehr schwierig ein absolut reines Transliterationssystem zu entwickeln. Daher weisen die meisten Transliterationssysteme einige Elemente der Transkription auf.¹²

2.1.2 Transliterationssysteme für die russische Schrift

2.1.2.1 Wichtige Merkmale der russischen Sprache und Schrift

Für ein besseres Verständnis der Fragestellungen, die mit der Transliteration der russischen Schrift in die lateinische verbunden sind, sollen zunächst ein paar grundlegende Informationen über die russische Sprache und Schrift aufgeführt werden.

Wie die meisten europäischen Sprachen gehört Russisch zur indoeuropäischen Sprachfamilie, und zwar in die Gruppe der slawischen Sprachen. Die slawischen Sprachen werden dabei in drei Untergruppen aufgeteilt: ostslawische (Russisch, Ukrainisch, Weißrussisch, auch Belarussisch genannt), westslawische (Polnisch, Tschechisch, Slowakisch) und südslawische Sprachen (Slowenisch, Bosnisch, Kroatisch, Serbisch, Mazedonisch,

⁸ Vgl. Wellisch, Hans H. (1978): The conversion of scripts - its nature, history, and utilization, S. 317.

⁹ Vgl. Duden (o. J.): Umschrift.

¹⁰ Vgl. Wiktionary (2018): Umschrift.

¹¹ Vgl. Wellisch, Hans H. (1978): The conversion of scripts - its nature, history, and utilization, S. 33; und Mangold, Max (1965): Transliteration und Transkription, S. 16.

¹² Vgl. ebd., S. 19.

Bulgarisch). Neben diesen zwölf großen Sprachen existiert eine Reihe kleinerer slawischer Sprachen.¹³

Aus historischen Gründen verfügen die slawischen Sprachen über unterschiedliche Schriftsysteme. In westslawischen Sprachen, deren Regionen von der katholischen Kirche geprägt waren, setzte sich die lateinische Schrift durch. Ostslawische Sprachen übernahmen zusammen mit dem orthodoxen Glauben das kyrillische Alphabet. Südslawische Sprachen werden teilweise lateinisch, teilweise kyrillisch geschrieben. Im Serbischen ist sogar die Verwendung beider Schriften möglich.¹⁴

Im Vergleich zu den anderen slawischen Sprachen verfügt Russisch über die größte Sprecherzahl. Weltweit ist Russisch an zehnter Stelle der am meisten gesprochenen Sprachen.¹⁵

Das russische Alphabet hat insgesamt 33 Buchstaben, bestehend aus 21 Konsonanten, 10 Vokalen, und dem sog. weichen („ь“) sowie harten Zeichen („Ъ“) ohne eigenen Lautwert. Nach der Oktoberrevolution fand im Jahr 1918 eine große Reform der russischen Sprache statt, die jedoch schon davor von der kaiserlichen Regierung veranlasst wurde. Im Rahmen der Reform wurden unter anderem aus dem russischen Alphabet die Buchstaben „ѣ“, „і“, „ѳ“ und „ѵ“ entfernt. Das harte Zeichen wurde zuerst am Wortende abgeschafft und im Wortinneren durch einen Bindestrich ersetzt, später jedoch wieder im Wortinneren eingeführt.¹⁶ Nach der Revolution gab es sogar Überlegungen, von der kyrillischen Schrift zur lateinischen zu wechseln. Die Idee wurde jedoch verworfen.¹⁷

An dieser Stelle soll auch darauf hingewiesen werden, dass die kyrillische Schrift nicht nur von slawischen Sprachen verwendet wird. Kyrillisch wurde bereits in der Zarenzeit von einigen nicht-slawischen Sprachen des Russischen Kaiserreiches geschrieben. In der Zeit der Sowjetunion fand jedoch eine größere Kyrillisierung der in der UdSSR gesprochenen nicht-slawischen Sprachen statt und das kyrillische Alphabet verbreitete sich unter anderem auf einige Turksprachen des Zentralen Asiens sowie iranische Sprachen wie Ossetisch und Tadschikisch. Das kyrillisierte Alphabet dieser Sprachen enthält in der Regel mehr Buchstaben als das russische kyrillische Alphabet.¹⁸ Da die Transliteration dieser nicht-slawischer Sprachen darüber hinaus einige Besonderheiten aufweist, wird sie gesondert in eigens dafür erlassenen Standards behandelt¹⁹ und soll in dieser Bachelor-Arbeit nur am Rande erwähnt werden.

¹³ Vgl. Heinz, Christof; Kuße, Holger, Hrsg. (2015): Slawischer Sprachvergleich für die Praxis, S. 63-65.

¹⁴ Vgl. ebd., S. 65.

¹⁵ Vgl. Rehder, Peter, Hrsg. (2009): Einführung in die slavischen Sprachen, S. 49.

¹⁶ Vgl. ebd., S. 50 und 54.

¹⁷ Vgl. Kempgen, Sebastian, Hrsg. (2009): Die slavischen Sprachen, S. 6

¹⁸ Vgl. Wellisch, Hans H. (1978): The conversion of scripts - its nature, history, and utilization, S. 107-108.

¹⁹ Wie DIN 1460:2:2010-3. Dazu vgl. Albrecht, Rita; Wiesner, Hans-Jörg, Hrsg. (2011): Umschriften in Bibliotheken, S. 17-45.

2.1.2.2 Transliterationsstandards

In einer der beiden Definitionen des Begriffes Transliteration wurde bereits die Existenz von mehreren internationalen und nationalen Transliterationsstandards angesprochen. Dies gilt in vollem Maße für die kyrillische Schrift, darunter für die russische. In der Mitte des 20. Jahrhunderts existierte in praktisch jeder europäischen Sprache ein eigenes Transliterationssystem für die kyrillische Schrift. In englischsprachigen Ländern wurden gleichzeitig mehr als ein Dutzend Transliterationssysteme genutzt. Die große Zahl von Standards sorgte für Verwirrung, die jedoch zumindest teilweise durch die Einführung des Standards ISO/R 9 im Jahr 1954 eliminiert wurde.²⁰

In Tabelle 1 werden ausgewählte derzeit geltende Transliterationsstandards verglichen und die Unterschiede zwischen ihnen veranschaulicht. Dabei werden die folgenden Standards berücksichtigt:

- der internationale Standard ISO 9:1995 "Information and documentation - Transliteration of Cyrillic characters into Latin characters - Slavic and non-Slavic languages"²¹ (ISO),
- die deutsche Norm DIN 1460 „Umschrift kyrillischer Alphabete slawischer Sprachen“²² (DIN),
- der russische Standard GOST 7.79-2000 (ISO 9-95) „Правила транслитерации кирилловского письма латинским алфавитом“ („Pravila transliteracii kirillovskogo pis'ma latinskim alfavitom")²³ mit Teil A, der identisch zu ISO 9:1995 ist, und Teil B, der Digrafen anstatt diakritischer Zeichen vorsieht, (GOST A und GOST B) sowie
- der Transliterationsstandard der Library of Congress und der American Library Association²⁴ (ALA-LC).

²⁰ Vgl. Wellisch, Hans H. (1978): The conversion of scripts - its nature, history, and utilization, S. 257.

²¹ Vgl. ISO (1995): Information and documentation - transliteration of Cyrillic characters into Latin characters - Slavic and non-Slavic languages, S. 3-6.

²² Vgl. Albrecht, Rita; Wiesner, Hans-Jörg, Hrsg. (2011): Umschriften in Bibliotheken, S. 14-16.

²³ Vgl. Mežgosudarstvennyj sovet po standartizacii, metrologii i sertifikacii (2000): GOST 7.79-2000 (ISO 9-95). Pravila transliteracii kirillovskogo pis'ma latinskim alfavitom, S. 4-7, 12-13.

²⁴ Vgl. Library of Congress (2012): ALA-LC Romanization Tables.

Tabelle 1. Vergleich ausgewählter Transliterationsstandards

Russisches Alphabet		Transliteration				
		ISO	DIN	GOST A	GOST B	ALA-LC
А	а	a				
Б	б	b				
В	в	v				
Г	г	g				
Д	д	d				
Е	е	e				
Ё	ё	ë	ë	ë	yo	ë
Ж	ж	ž	ž	ž	zh	zh
З	з	z				
И	и	i				
Й	й	j	j	j	j	ï
К	к	k				
Л	л	l				
М	м	m				
Н	н	n				
О	о	o				
П	п	p				
Р	р	r				
С	с	s				
Т	т	t				
У	у	u				
Ф	ф	f				
Х	х	h	ch	h	x	kh
Ц	ц	c	c	c	cz, c	ts
Ч	ч	č	č	č	ch	ch
Ш	ш	š	š	š	sh	sh
Щ	щ	ŝ	šč	ŝ	shh	shch
Ъ	ъ	"				

Russisches Alphabet		Transliteration				
		ISO	DIN	GOST A	GOST B	ALA-LC
Ы	ы	y	y	y	y`	y
Ь	ь	'				
Э	э	è	é	è	e`	é
Ю	ю	û	ju	û	yu	îu
Я	я	â	ja	â	ya	îa

Tabelle 1 zeigt sehr deutlich, wie unterschiedlich die verglichenen Transliterationsstandards sind. Während sie sich in einem Teil der russisch-kyrillischen Buchstaben, für die es einfache lateinische Äquivalente gibt, durchaus einig sind, werden sehr unterschiedliche Lösungen in Bezug auf Buchstaben ohne eindeutige lateinische Äquivalente angeboten. Dies betrifft die russischen Buchstaben „ё“, „ж“, „й“, „х“, „ц“, „ч“, „щ“, „ш“, „ы“, „э“, „ю“ und „я“. Sie werden in der lateinischen Schrift entweder durch die Verwendung der diakritischen Zeichen oder durch eine Kombination von zwei lateinischen Buchstaben (manchmal zusätzlich mit diakritischen Zeichen) wiedergegeben.

2.1.3 Regelungen für bibliothekarische Transliteration in Deutschland

Das erste allgemein verbreitete deutschsprachige Katalogisierungsregelwerk erschien 1899 unter dem Titel „Instruktionen für die alphabetischen Kataloge der preußischen Bibliotheken“, kurz Preußische Instruktionen genannt.²⁵ Gemäß Punkt 1 und 2 § 4 der Preußischen Instruktionen war der Titel in lateinischer Schrift aufzunehmen. Jede andere Schriftart außer der Griechischen war nach einem in Anlage II aufgeführten Schema zu „transkribieren“.²⁶ Damals wurde der Begriff Transkription verwendet, obwohl es sich hier eher um Transliteration handelte, da dem Vorgang nicht die phonetische, sondern die schriftliche Form zugrunde lag. Das Transliterationsschema der Preußischen Instruktionen für die russische Schrift wird in der Tabelle 2 im Vergleich mit anderen in diesem Unterpunkt erwähnten Transliterationsschemas aufgeführt.

Im Ergebnis der Bestrebungen nach einer internationalen Annäherung bibliothekarischer Regelwerke wurden die Preußischen Instruktionen durch die 1977 erschienenen „Regeln für die alphabetische Katalogisierung in Bibliotheken“ (kurz RAK) abgelöst.²⁷ Laut § 116 der RAK war die Einheitsaufnahme in lateinischer Schrift unabhängig von der

²⁵ Vgl. Gantert, Klaus (2016): Bibliothekarisches Grundwissen, S. 180.

²⁶ Vgl. o. A. (1966): Instruktionen für die alphabetischen Kataloge der preussischen Bibliotheken vom 10. Mai 1899, S. 5.

²⁷ Vgl. Gantert, Klaus (2016): Bibliothekarisches Grundwissen, S. 181.

Schrift der Vorlage zu erfassen. Nichtlateinische Schriftarten sollten gemäß Transliterationstabellen in Anlage 5 umgeschrieben werden.²⁸

Ein Teil der Anlage 5, und zwar die „Tabellen für die Umschrift nichtlateinischer Schriftzeichen in die Buchstaben der lateinischen Schrift“ wurde als Vorabdruck bereits 1975 veröffentlicht und lieferte interessante Einblicke in die Überlegungen, die zur Entstehung dieser Tabelle geführt haben.²⁹ Zu dem Zeitpunkt haben die meisten wissenschaftlichen Bibliotheken in Deutschland für die Transliteration der slawischen kyrillischen Buchstaben die Norm DIN 1460 in der damaligen Fassung aus dem Jahr 1962 verwendet.³⁰ Die im Jahr 1973 durchgeführte Umfrage der Anwender zeigte, dass die Befragten in der Auswahl zwischen DIN 1460, ISO-R 9 Grundtabelle und ISO-R 9 mit allen Varianten sich fast einstimmig für die Beibehaltung der DIN 1460 ausgesprochen haben, die als Folge in die Anlage 5 zu den RAK aufgenommen wurde.³¹

Im Rahmen der weiteren Entwicklungen erschien im Jahr 1982 die neue Fassung der DIN 1460.³² Die einheitlichen RAK wurden in die Regeln für wissenschaftliche Bibliotheken (RAK-WB, erschienen 1983) und die Regeln für öffentliche Bibliotheken (RAK-ÖB, erschienen 1986) aufgeteilt.³³ Die Anwendung der Transliterationsregeln gemäß DIN 1460 blieb weiterhin beibehalten.³⁴

2015 wurden die RAK im deutschsprachigen Raum durch das internationale Regelwerk „Resource Description and Access“ (RDA) ersetzt.³⁵ Punkt 5.4. RDA „Sprache und Schrift“ sieht folgende Regelungen in Bezug auf fremdsprachige Werke vor:

„Erfassen Sie Titel für Werke in der Sprache und Schrift, in der sie in den Quellen erscheinen, aus denen sie entnommen sind.“

Alternative: Erfassen Sie eine transliterierte Form des Titels entweder als Ersatz für die Form, die in der Quelle erscheint oder zusätzlich.

Erfassen Sie Merkmale eines Werks oder einer Expression in der Sprache und Schrift, die die Agentur bevorzugt, welche die Daten erstellt, außer wenn die Bestimmungen für ein bestimmtes Element etwas anderes vorgeben.“³⁶

²⁸ Vgl. o. A. (1977): Regeln für die alphabetische Katalogisierung, S. 15.

²⁹ Vgl. o. A. (1975): Regeln für die alphabetische Katalogisierung. Anlage 5. Tabellen für die Umschrift nichtlateinischer Schriftzeichen in die Buchstaben der lateinischen Schrift, S. II-IV.

³⁰ Vgl. DIN (1962): DIN 1460 Transliteration slawischer kyrillischer Buchstaben.

³¹ Vgl. o. A. (1975): Regeln für die alphabetische Katalogisierung. Anlage 5. Tabellen für die Umschrift nichtlateinischer Schriftzeichen in die Buchstaben der lateinischen Schrift, S. III-IV.

³² Vgl. Albrecht, Rita; Wiesner, Hans-Jörg, Hrsg. (2011): Umschriften in Bibliotheken. Anwendung, Verfahren, Tabellen, S. 13-16.

³³ Vgl. Gantert, Klaus (2016): Bibliothekarisches Grundwissen, S. 181.

³⁴ Vgl. Deutsche Nationalbibliothek (2007): Regeln für die alphabetische Katalogisierung in wissenschaftlichen Bibliotheken. S. 427-428.

³⁵ Vgl. Gantert, Klaus (2016): Bibliothekarisches Grundwissen, S. 187.

³⁶ RDA Toolkit, Zugriff am 26.11.2019.

Die zugehörigen Anwendungsrichtlinien D-A-CH sehen an dieser Stelle die Anwendung der Alternative an, somit soll der Titel der Werke transliteriert werden. Im Reiter „Werkzeuge“ des Toolkits stehen unter anderem Transliterationstabellen für den deutschsprachigen Raum zur Verfügung, darunter die DIN 1460 „Umschrift kyrillischer Alphabete slawischer Sprachen“ in der Fassung von April 1982. Somit bleiben die Transliterationsregeln für das russische Alphabet auch nach der Einführung des neuen Regelwerks unverändert.

Momentan wird die derzeit geltende Fassung der DIN 1460 aktualisiert. Der bereits zur Diskussion veröffentlichte Entwurf sieht jedoch keine Änderungen in der Umschrifttabelle für die russische Sprache vor.³⁷

Wie bereits oben angekündigt, werden die angesprochenen deutschen bibliothekarischen Transliterationsschemas in der folgenden Tabelle 2 verglichen, und zwar:

- die Preußischen Instruktionen³⁸ (PI),
- die Regeln für die alphabetische Katalogisierung³⁹ (RAK), basierend auf DIN 1460 in der Fassung von Oktober 1962,
- sowie die aktuell geltende DIN 1460 „Umschrift kyrillischer Alphabete slawischer Sprachen“⁴⁰ in der Fassung von 1982 (DIN).

Tabelle 2. Vergleich deutscher bibliothekarischer Transliterationsschemas

Russisches Alphabet		Transliterationsschema		
		PI	RAK	DIN
A	а	a		
Б	б	b		
В	в	v		
Г	г	g		
Д	д	d		
Е	е	e		
Ё	ё	---	ë	
Ж	ж	ž		
З	з	z		

³⁷ Vgl. DIN (2019): Norm-Entwurf DIN 1460. Umschrift kyrillischer Alphabete - Umschrift kyrillischer Alphabete slawischer Sprachen.

³⁸ Vgl. o. A. (1966): Instruktionen für die alphabetischen Kataloge der preussischen Bibliotheken vom 10. Mai 1899, S. 142.

³⁹ Vgl. o. A. (1977): Regeln für die alphabetische Katalogisierung, S. 378-379.

⁴⁰ Vgl. Albrecht, Rita; Wiesner, Hans-Jörg, Hrsg. (2011): Umschriften in Bibliotheken. Anwendung, Verfahren, Tabellen, S. 13-16.

Russisches Alphabet		Transliterationsschema		
		PI	RAK	DIN
И	и	i		
Й	й	j		
К	к	k		
Л	л	l		
М	м	m		
Н	н	n		
О	о	o		
П	п	p		
Р	р	r		
С	с	s		
Т	т	t		
У	у	u		
Ф	ф	f		
Х	х	ch		
Ц	ц	c		
Ч	ч	č		
Ш	ш	š		
Щ	щ	šč		
Ъ	ъ	1)	" 3)	am Wortende (") ⁴⁾ im Wortinneren "
Ы	ы	y		
Ь	ь	' 2)	'	' 5)
Э	э	ě		
Ю	ю	ju		
Я	я	ja		

1) wird am Wortende nicht berücksichtigt, im Wortinneren aber durch Bindestrich wiedergeben.

2) bei Oberlängen neben, sonst über dem vorangehenden Buchstaben.

3) im Wortinneren, kann am Ende entfallen.

4) Nur noch im Wortinneren gebräuchlich und in dieser Stellung entsprechend seiner Funktion als Trennungszeichen transliteriert. Am Wortende hatte „б“ bis zur Schriftreform 1918 die Funktion eines Härtezeichens. „ zu schreiben als Doppelapostroph bzw. Anführungszeichen.

5) ' zu schreiben als Apostroph.

Im Gegensatz zur Tabelle 1, die internationale Transliterationsstandards vergleicht, zeigt Tabelle 2 deutlich weniger Unterschiede zwischen den einzelnen Transliterationsschemas. Die Preußischen Instruktionen sahen kein lateinisches Äquivalent für den russischen Buchstaben „ё“ vor. Zusätzlich gab es Variationen in der Behandlung der Buchstaben „ѣ“ und „ѥ“, die in den Transliterationsschemas in Fußnoten erläutert wurde. Diese Unterschiede können allerdings als unwesentlich betrachtet werden. Somit verfügt das deutschsprachige Bibliothekswesen in Bezug auf das russische Alphabet über eine langjährige Transliterationstradition, die sich in den letzten hundert Jahren kaum geändert hat.

2.1.4 Regelungen für die Erfassung nicht-lateinischer Schriften in K10plus

In der Verbunddatenbank K10plus wird die Erfassung nicht-lateinischer Schriften im gleichnamigen Katalogisierungshandbuch geregelt. Aus der Einleitung lässt sich die folgende Grundregel formulieren:

„Im K10plus besteht die Möglichkeit, bestimmte Felder der bibliografischen Beschreibung und der Normdaten sowohl in lateinischer als auch in nicht-lateinischer Schrift zu erfassen. Dabei ist die nicht-lateinische Schrift nur zusätzlich zur lateinischen Schrift zugelassen. [...] In welchem Umfang Felder mit nicht-lateinischen Zeichen erfasst werden, liegt im Ermessen der jeweiligen Bibliothek.“⁴¹

Im Punkt 4 des Handbuches werden die zu verwendenden Transliterationstabellen aufgelistet, darunter die DIN 1460 in der Fassung 1982 für das kyrillische Alphabet der slawischen Sprachen.⁴² Außerdem steht in WinIBW ein automatisches Transliterationstool zur Verfügung. Das Tool ist eher als Hilfe gedacht, denn die Ergebnisse der automatischen Transliteration sollten überprüft werden.⁴³ Eingeführt wurde das Tool im Jahr 2013.⁴⁴

2.1.5 Kritik an Transliteration in bibliothekarischen Datenbanken

In vorangehenden Punkten 2.1.3 und 2.1.4 wurde beschrieben, wie die Erfassung nicht-lateinischer Schriften in deutschsprachigen Regelwerken behandelt wurde bzw. wird.

⁴¹ SWB-GBW (2019): Erfassung nicht-lateinischer Schriften, S. 2.

⁴² Vgl. ebd., S. 9.

⁴³ Vgl. ebd., S. 2.

⁴⁴ Vgl. Scherer, Birgit (2013): Neue WinIBW 3.6.2.1, E-Mail an BSZ-Forum vom 13.09.2013. URL: https://wiki.bsz-bw.de/lib/exe/fetch.php?media=v-team:mailwinibw_3.6.2.1.pdf (26.11.2019)

Die Hauptregel ist dabei die Transliteration nicht-lateinischer Titel gemäß vorgegebener Transliterationstabellen. Erst seit der Einführung von RDA ist die Erfassung in der Originalschrift zusätzlich zur transliterierten Variante möglich. Auch im englischsprachigen Raum war Transliteration lange die Hauptregel.⁴⁵

Die lange Tradition, in der die Transliteration nicht-lateinischer Schriften als die einzige Möglichkeit bei der Erfassung entsprechender Werke galt, hatte ihre Gründe. In der Zeit der Zettelkataloge und Schreibmaschinen war die Anzahl der verfügbaren Schriftzeichen technisch bedingt beschränkt und es wäre nicht möglich gewesen, Titel in unterschiedlichen Sprachen in Originalschrift zu erfassen. In der früheren Zeit des Internets gab es Probleme bei der Anzeige nicht-lateinischer Schriften wegen der Verwendung unterschiedlicher Schriftzeichen-Kodierungen, sodass nicht-lateinische Wörter oft als eine Reihe von Fragezeichen oder anderer sinnloser Platzhalter angezeigt wurden.⁴⁶

Erst die weitere technische Entwicklung ermöglichte die Verwendung der Originalschrift in OPACs. Im Südwestdeutschen Bibliotheksverbund (SWB) wurde die Erfassung in Originalschrift wegen des entsprechenden Bedarfs im Februar 2009 noch lange vor der Einführung von RDA ermöglicht.⁴⁷

Transliteration als der einzige mögliche Ansatz bei der Erfassung nicht-lateinischer Titel war dabei bei Weitem nicht unumstritten und stieß in der bibliothekarischen Community teilweise auf Kritik.⁴⁸ Um nach Werken in nicht-lateinischen Schriften zu suchen, müssen die Nutzer*innen die Regeln kennen, gemäß welchen diese Werke in bibliothekarischen Katalogen erfasst worden sind. Viele Nutzer*innen, die gewohnt sind, die Werke in nicht-lateinischen Schriften zu lesen, gehen jedoch eher von ihrer Erfassung in Originalschrift aus. Außerdem entsteht dabei eine Ungleichheit beim Zugang zum Bibliotheksbestand, weil solche Nutzer*innen mehr Hürden beim Zugang zu Bibliotheksressourcen haben.⁴⁹

Erfreulicherweise ist diese Kritik zumindest in Bezug auf die in dieser Bachelor-Arbeit untersuchte Datenbank K10plus nicht mehr aktuell, denn die Erfassung der Titel in Originalschrift ist derzeit sowohl technisch möglich als auch durch die entsprechenden Regelwerke erlaubt.

Nachdem in diesem Punkt die Hintergründe zur Transliteration beschrieben und die Begriffe geklärt sowie die internationalen und nationalen Transliterationsstandards erläutert wurden, soll der nächste Punkt dieses Kapitels Tippfehlern in bibliothekarischen Datenbanken und ihrer Behandlung im bibliothekarischen Bereich sowie dem Stand der Forschung zu diesem Thema gewidmet werden.

⁴⁵ Vgl. Agenbroad, James E. (2006): Romanization Is Not Enough, S. 22-25.

⁴⁶ Vgl. Brewer, Michael M. (2009): Romanization of Cyrillic Script: Core Competencies and Basic Research Strategies for Slavic Students, Scholars, and Educators, S. 245.

⁴⁷ Vgl. Horny, Silke; Böhner, Susanne (2009): Erfassung originalsprachlicher Titel, E-Mail an BSZ-Forum vom 04.02.2009.

⁴⁸ Vgl. Agenbroad, James E. (2006): Romanization Is Not Enough, S. 25-26.

⁴⁹ Vgl. ebd., S. 22 und 24.

2.2 OPAC-Qualität und Tippfehler

2.2.1 Fehlerfreier OPAC als Qualitätskriterium

Als Erstes soll im Rahmen dieses Punktes ermittelt werden, ob ein fehlerfreier OPAC überhaupt als ein Qualitätskriterium für Bibliotheken gilt. Der gesunde Menschenverstand lässt logischerweise annehmen, dass korrekte Datensätze für einen gut funktionierenden OPAC wichtig sein müssten. Da eigene Annahmen sich als trügerisch erweisen können, soll überprüft werden, ob die aktuelle bibliothekarische Literatur von demselben Standpunkt ausgeht.

Laut Matthews gehört Katalogisierung zu den technischen Dienstleistungen einer Bibliothek und wird somit als diese auch evaluiert.⁵⁰ Tippfehler in bibliothekarischen Datenbanken wirken sich negativ auf die Qualität der technischen Dienstleistungen aus, da dadurch das Auffinden von notwendigen Informationen erschwert wird. Tippfehler in den primären Feldern wie Autor, Titel und Schlagwörter sind besonders problematisch. Matthews schlägt Bibliotheken vor, die Qualität der eigenen OPACs zu prüfen, indem sie eine Stichprobe von physischen Medien und die entsprechenden bibliografischen Datensätze vergleichen.⁵¹

Poll beschreibt unter Leistungskennzahlen, mit welchen die Leistungen einer Bibliothek gemessen werden können, den Indikator „Verfügbarkeit der benötigten Titel“.⁵² Dieser Indikator ist Teil des Aspektes „Ressourcen und Infrastruktur“ und seiner Untergruppe „Bestand“ und beurteilt die Befriedigung des Informationsbedarfs der Nutzer*innen. Gemessen wird dabei die Prozentzahl der durch die Nutzer*innen benötigten und dabei direkt verfügbaren Titel. Situationen, wenn ein verfügbarer Titel wegen fehlerhaft erfassten (darunter Tippfehler enthaltenden) Datensätze im Katalog nicht gefunden werden kann, wirken sich dementsprechend negativ auf die Qualität der Dienstleistungen einer Bibliothek aus.⁵³

Auch Klauß weist auf negative Folgen hin, die in Situationen entstehen, wenn ein gesuchter und verfügbarer Titel trotz sorgfältiger Recherche nicht gefunden werden kann, weil der Datensatz Tippfehler enthält. Diese Schäden können unnötige Kosten verursachen, indem ein Titel neu beschaffen oder per Fernleihe bestellt wird oder durch die Nutzer*innen in einer anderen Bibliothek vor Ort besorgt werden muss. Auch wenn keine direkten finanziellen Kosten entstehen, entsteht ein Schaden im kulturellen oder wissenschaftlichen Sinne, indem ein Titel nicht verwendet werden kann, der in der jeweiligen Situation oder für das angestrebte Forschungsprojekt wertvoll gewesen wäre.⁵⁴

⁵⁰ Vgl. Matthews, Joseph R.; Hinchliffe, Lisa (2018): The evaluation and measurement of library services, S. 103.

⁵¹ Vgl. ebd., S. 113.

⁵² Vgl. Poll, Roswitha; Boekhorst, Peter te (2007): Measuring quality. Performance measurement in libraries, S. 43-44.

⁵³ Vgl. ebd., S. 64-64.

⁵⁴ Vgl. Klauß, Henning (2001): Tippfehler in Bibliothekskatalogen. Möglichkeiten einer EDV-gestützten Ermittlung, S. 868.

Somit wird bestätigt, dass korrekte Erfassung der bibliografischen Informationen in bibliothekarischen Datenbanken eine wichtige Rolle in der Arbeit einer Bibliothek spielt. Tippfehler können sich dabei negativ auf die Qualität der grundlegenden Dienstleistungen einer Bibliothek auswirken und sowohl direkte Kosten für die Bibliothek und ihre Nutzer*innen als auch unmittelbare Schäden verursachen.

2.2.2 Tippfehler in bibliothekarischen Datenbanken

Die Diskussion über die Tippfehler in bibliothekarischen Datenbanken ist bei Weitem nicht neu. Das Thema kam bereits Ende der siebziger Jahre auf.

Bourne analysierte 1977 die Häufigkeit von Tippfehlern im Index in elf unterschiedlichen Datenbanken. Die Zahl an fehlerhaften Wörtern variierte von 0,47 % bis 22,8 % der Gesamtwörter der Stichprobe aus dem jeweiligen Index und betrug im Durchschnitt 10,8 %. Um die tatsächliche Auswirkung dieser fehlerhaften Indexeinträge zu prüfen, analysierte Bourne die Anzahl der Datensätze, die mit einem solchen Indexeintrag verknüpft und somit direkt vom Tippfehler betroffen waren. Die Anzahl solcher Datensätze lag dabei zwischen 0,01 % und 0,63 % von der Gesamtzahl der Datensätze aus der Stichprobe des jeweiligen Index, im Durchschnitt bei 0,24 % der Gesamtzahl der Datensätze in der Gesamtstichprobe.⁵⁵

Ryans untersuchte OCLC-Datensätze auf Fehler aller Art und ihre Verteilung auf unterschiedliche Fehlerarten. In den 700 geprüften Datensätzen wurden 393 einzelne Fehler in insgesamt 283 Datensätzen gefunden. Bei vielen davon handelte es sich um Rechtschreib- oder Tippfehler. Da Ryans diese als unbedeutende Fehler betrachtete, wurden sie nicht weiter analysiert. Die genaue Anzahl wurde ebenfalls nicht genannt.⁵⁶

1991 empfahl Jeffrey Beall der bibliothekarischen Community, eigene OPACs auf Tippfehler zu testen. Er stellte eine Liste von zehn oft vorkommenden Wörtern zusammen, die Tippfehler enthielten, und schlug vor, nach diesen im eigenen OPAC zu recherchieren. Auf Grundlage der Ergebnisse sollte ein Wert kalkuliert werden, der einen Vergleich zwischen Datenbanken ähnlicher Größe erlauben würde. Diesen Test nannte Beall „Dirty Database Test“.⁵⁷

Bealls kurze, nur wenige Absätze lange Beitrag erregte große Aufmerksamkeit. Nachdem im Ergebnis des „Dirty Database Tests“ an der Adelphi University in New York zwei der zehn vorgeschlagenen Tippfehler festgestellt worden sind, überprüfte Terry Ballard den ganzen Index, der damals aus 117.000 Wörtern bestand, und stellte eine umfassende Liste der gefundenen Tippfehler zusammen. Ballard kam zum Ergebnis, dass Tippfehler öfter in sehr verbreiteten Wörtern vorkommen, die aus acht Buchstaben und

⁵⁵ Vgl. Bourne, Charles P. (1977): Frequency and impact of spelling errors in bibliographic data bases, S. 1–12.

⁵⁶ Vgl. Ryans, Cynthia C. (1978): A study of errors found in non-MARC cataloging in a machine-assisted system, S. 125–132.

⁵⁷ Vgl. Beall, Jeffrey (1991): The Dirty Database Test, S. 197.

drei oder mehr Silben bestehen.⁵⁸ Die meisten Tippfehler kamen im Titelfeld vor, in 40 % der Fälle waren dabei die ersten drei Titelwörter betroffen. Ballard kam zum Schluss, dass die von ihm festgestellten Fehler universaler Natur waren und unterschiedliche Datenbanken in unterschiedlichen Ländern betrafen. Er veröffentlichte die gesamte Liste der ermittelten Tippfehler, die als Grundlage für die Überprüfung anderer Datenbanken genutzt werden konnte.⁵⁹ Mit der Zeit wurde aus der Liste ein kooperatives Projekt, das außer der Erweiterung der gelisteten Tippfehler in Form eines Wikis das Blog „Typo of the Day“ umschloss.⁶⁰ Die Liste wird allerdings seit 2016⁶¹ und das Blog seit 2017 nicht aktualisiert.⁶²

Auch Beall selbst beschäftigte sich weiter mit Tippfehlern in bibliothekarischen Katalogen. Er verwendete die erwähnte Liste von Ballard, um Tippfehler in Normdatensätzen ausfindig zu machen und schlug vor, für diese eine gesonderte Liste einzuführen.⁶³

Außerdem untersuchten Beall und Kafadar, wie viele Tippfehler bei der Übernahme der Datensätze korrigiert werden. Dazu wählten sie zufällig hundert Tippfehler aus Ballards Liste aus und suchten in der bibliografischen Datenbank von OCLC Datensätze aus, die diese Tippfehler enthielten (sog. Master-Records). Danach prüften sie in den Katalogen von fünf zufällig ausgewählten Bibliotheken, die diese Master-Records in ihre lokalen Kataloge übernommen haben, ob der Tippfehler während der Datenübernahme korrigiert worden ist. Dabei wurde sorgfältig überprüft, ob es sich wirklich um denselben Master-Record handelte. Die Ergebnisse der Studie waren ernüchternd – bei der Datenübernahme wurden lediglich 35,8 % der Tippfehler berichtigt, 64,2 % blieben unkorrigiert.⁶⁴

Dieselben Autoren untersuchten die Auswirkung der Tippfehler auf das Retrieval am Beispiel von Worldcat. Sie nahmen an, dass Tippfehler einen negativen Einfluss auf das Retrieval haben, außer in Fällen, wenn das falsch geschriebene Wort auch an einer anderen Stelle im Datensatz vorkommt und dort richtig geschrieben ist. Autoren wählten 135 Wörter mit Tippfehlern aus, die sie nach allgemeiner Worthäufigkeit sowie Tippfehlerhäufigkeit gemäß Ballard sortierten. Nach diesen fehlerhaft geschriebenen Wörtern wurde in trunkierter Form mittels allgemeiner Stichwortsuche (keyword search) im Worldcat zwei Mal recherchiert. Bei der ersten Recherche wurde die Zahl der Datensätze ermittelt, die das Wort mit fehlerhafter Schreibweise allgemein enthielten (z. B. Recherche nach „congress*“ ergab 282 Treffer). Bei der zweiten Recherche wurde gezielt nach

⁵⁸ Eines der öfter falsch geschriebenen Wörter war dabei „commercial“ in der Form von „commerical“. Die in K10plus am 19.11.2019 durchgeführte Titelstichwort-Suche ergab 171 Treffer für „commerical“ im Vergleich zu Ballards 8 Treffern, was vermutlich auf den Größenunterschied der Datenbanken zurückzuführen ist.

⁵⁹ Vgl. Ballard, Terry (1992): Spelling and Typographical Errors in Library Databases, S. 14–19.

⁶⁰ Vgl. Ballard, Terry (2008): Systematic Identification of Typographical Errors in Library Catalogs, S. 30–32.

⁶¹ Vgl. o. A. (2016): Typo of the day for libraries – Index.

⁶² Vgl. o. A. (2017): Typo of the day for libraries.

⁶³ Vgl. Beall, Jeffrey (2004): Using OCLC Connexion to find typographical errors in authority records, S. 71–75.

⁶⁴ Vgl. Beall, Jeffrey; Kafadar, Karen (2004): The Effectiveness of Copy Cataloging at Eliminating Typographical Errors in Shared Bibliographic Records, S. 92–96.

Datensätzen recherchiert, die das Wort mit fehlerhafter Schreibweise enthielten und dieses Wort an keiner anderen Stelle im Datensatz in seiner richtigen Schreibweise vorkam (z. B. Recherche nach „congress* not congress“ ergab 246 Treffer). In solchen Datensätzen hätte man dieses Wort nur dann finden können, wenn man gezielt nach der falschen Schreibweise recherchiert. Durch eine korrekte Anfrage ohne Tippfehler waren diese Datensätze nicht auffindbar. Im Beispielsfall „congress“ wären es 246 Datensätze von insgesamt 282, was einem Anteil von 87,2 % entspricht. 32 Datensätze (12,8 %) waren dabei auch mit korrekter Anfrage mittels allgemeine Stichwortsuche auffindbar, obwohl sie einen Tippfehler enthielten.⁶⁵

Der Anteil der durch korrekte Anfrage nicht auffindbarer Datensätze variierte in der Stichprobe je nach Wort zwischen 40 % bis 97 %. Entscheidend dafür war vor allem die Worthäufigkeit – je häufiger ein Wort in bibliothekarischen Datenbanken vorkommt, desto höher ist die Wahrscheinlichkeit, dass es an einer anderen Stelle im Datensatz richtig geschrieben wird und der negative Einfluss des Tippfehlers auf das Retrieval somit niedriger ausfällt. Bei seltener vorkommenden Wörtern ist die Auswirkung der Tippfehler auf das Retrieval wiederum größer.⁶⁶

Diese Ergebnisse von Beall und Kafadar stehen im Widerspruch zu einer zeitlich vorangehenden Studie von Randall, die ebenfalls die Auswirkung der Tippfehler auf das Retrieval untersuchte und zum Schluss kam, dass die meisten Tippfehler nicht retrievalrelevant sind. In ihrer Studie teilte Randall Tippfehler in zwei Gruppen auf: einmalige (engl. „unique“) Tippfehler, wenn sie nur an einer Stelle im Datensatz vorkamen, und redundante (engl. „redundant“) Tippfehler, wenn das falsch geschriebene Wort zusätzlich an einer anderen Stelle im Datensatz vorkam und dabei korrekt geschrieben war. Im Ergebnis kam Randall zum Schluss, dass die meisten Tippfehler redundant waren und die Nutzer*innen nicht daran hinderten, die gewünschten Datensätze zu finden.⁶⁷

Die Untersuchung von Randall weist jedoch im Vergleich zur Studie von Beall und Kafadar einige Einschränkungen auf. Sie basiert auf der Überprüfung von lediglich drei Wörtern (economy, economic und economics), die zudem wurzelverwandt sind und mithilfe von Trunkierung mit einem Suchstring hätten abgefragt werden können. Die dadurch gezogene Stichprobe bestand aus 23, 30 und 40 Datensätzen in drei untersuchten Datenbanken und war somit eher gering. Außerdem beruhen die Schlussfolgerungen von Randall lediglich auf der Annahme, dass redundante Tippfehler keinen negativen Einfluss auf das Retrieval haben. Diese Annahme wurde jedoch weder in Bezug auf einen speziellen Index noch in Bezug auf den ALL-Index überprüft. Darüber hinaus ist die Anzahl der durch Randall festgestellten einmaligen und somit retrievalrelevanten Tippfehlern nicht so niedrig, dass sie als unwichtig eingestuft werden könnte, denn es waren je

⁶⁵ Vgl. Beall, Jeffrey; Kafadar, Karen (2007): Measuring Typographical Errors' Impact on Retrieval in Bibliographic Databases, S. 197–204.

⁶⁶ Vgl. ebd., S. 207.

⁶⁷ Vgl. Randall, Barbara Nichols (1999): Spelling Errors in the Database: Shadow or Substance?, S. 161–169.

nach untersuchter Datenbank 18 %, 40 % bzw. 43 % der Tippfehler enthaltenden Datensätze.⁶⁸

Im Gegensatz zu Randall wählten Beall und Kafadar stichprobenartig 135 unterschiedliche Wörter aus und verwendeten die Trunkierung, um unterschiedliche Wortformen zu berücksichtigen. Außerdem wurden im Rahmen der Studie für jedes der 135 Wörter tatsächliche Recherchen im Worldcat durchgeführt.⁶⁹ Somit wirken die Ergebnisse von Beall und Kafadar in einigen Punkten deutlich überzeugender.

Cherukodan et al. analysierten ebenfalls das Vorkommen von Tippfehlern in bibliothekarischen Datenbanken. Ihre Untersuchung basierte auf nur einem Wort (zehn falsch geschriebene Varianten des Wortes „management“) und sollte vor allem feststellen, ob das Problem von Tippfehlern für Bibliotheken noch aktuell ist. In der Studie wurden die OPACs von fünf Nationalbibliotheken (British Library, Library of Congress, National Library of India, National Library of Australia, National Library of South Africa), fünf Bibliotheken von führenden Universitäten (Stanford, Oxford, Harvard, Princeton und Massachusetts Institute of Technology) sowie vier große Online-Datenbanken (ScienceDirect, ProQuest ABI/INFORM Complete, Taylor & Francis und Emerald) untersucht. Nur der OPAC des California Institute of Technology enthielt keine Tippfehler im ausgesuchten Wort, lieferte somit keine Treffer und wurde daher in der Studie nicht weiter berücksichtigt. In allen anderen untersuchten OPACs und Datenbanken waren mehrere oder fast alle der zehn Tippfehlervariationen vertreten. Dies zeigt, dass kaum eine Bibliothek von Tippfehlern verschont bleibt, und zwar unabhängig von ihrem Typ, Größe oder Region.⁷⁰

Die in diesem Punkt aufgeführte Beschreibung der bereits existierenden Studien zum Thema Tippfehler in bibliothekarischen Datenbanken spiegelt den aktuellen Stand der Forschung wider, berücksichtigt jedoch nur ausgewählte Werke und hat somit keinen Anspruch auf Vollständigkeit, die im Rahmen einer Bachelor-Arbeit kaum gewährleistet werden kann. Außerdem stammen praktisch alle Studien aus dem anglo-amerikanischen Raum, der einige Unterschiede zum deutschen Verbundsystem aufweist. Daher können ihre Ergebnisse nur bedingt auf den deutschsprachigen Raum angewendet werden.

Alle oben beschriebenen Untersuchungen beschäftigten sich zudem mit allgemeinen Tippfehlern. Trotz umfassender Recherche konnten keine Werke gefunden werden, die gezielt nur Tippfehler analysieren, welche im Ergebnis von Transliteration entstehen. Dieses spezielle Thema scheint somit nicht ausreichend erforscht zu sein und bietet für die vorliegende Bachelor-Arbeit die Möglichkeit, etwas Neues zum aktuellen Forschungsstand beizutragen.

⁶⁸ Vgl. ebd., S. 163-165.

⁶⁹ Vgl. Beall, Jeffrey; Kafadar, Karen (2007): Measuring Typographical Errors' Impact on Retrieval in Bibliographic Databases, S. 199-203.

⁷⁰ Vgl. Cherukodan, Surendran et al. (2013): Analytical Study of Typographical Errors in OPACS and Corrective Measures, S. 116–123.

3 Analyse der Tippfehler in K10plus

3.1 Beschreibung des Forschungsvorhabens

3.1.1 Forschungsfragen

Die Analyse der Tippfehler in der Verbunddatenbank K10plus wurde mit dem Ziel durchgeführt, Antworten auf folgende Forschungsfragen zu finden:

- Welche Transliterationsfehler entstehen bei der Formalerschließung russischsprachiger Monografien? Lassen sich diese Fehler nach Fehlerarten kategorisieren?
- Wie viele Datensätze der Stichprobe (Prozentzahl) enthalten Fehler in den analysierten Feldern?
- Wie oft kommen einzelne Fehlerarten vor? Können dabei Muster identifiziert werden?

3.1.2 Methodik

Im Rahmen dieser Bachelor-Arbeit wurden in der Verbunddatenbank K10plus Datensätze analysiert, in welchen Russisch als Sprache angegeben war und die Felder mit transliterierten bibliografischen Informationen enthielten.

Es wurden Felder analysiert, bei denen der größte Transliterationsbedarf besteht und die sich ohne großen Zeitaufwand überprüfen lassen. Dies war notwendig, weil die Bearbeitungsfrist für diese Bachelor-Arbeit durch die Studien- und Prüfungsordnung eingeschränkt ist, bei der Analyse jedoch eine ausreichend große Stichprobe angestrebt wurde.

Der größte Transliterationsbedarf besteht in der Formalerschließung typischerweise beim Haupttitel bzw. Titelzusatz, der Verantwortlichkeitsangabe und bei den Namen der geistigen Schöpfer und sonstiger Personen und Körperschaften. Auf die Auswertung der Personen- und Körperschaftsnamen und somit auch der Verantwortlichkeitsangabe wurde verzichtet. Die Überprüfung der korrekten Transliteration ist in diesen Feldern nur eingeschränkt möglich – es können nur einige gängige Namen sowie Namen von bekannten Persönlichkeiten oder Körperschaften überprüft werden. In den meisten Fällen müsste der Name im Original überprüft werden, was einen sehr hohen Arbeitsaufwand mit sich ziehen würde. Aus diesem Grund wurde die Analyse nur auf den Haupttitel und Titelzusätze beschränkt, was dem Feld 4000 in K10plus (jedoch ohne Unterfeld Verantwortlichkeitsangabe) entspricht.

Um die Größe der Stichprobe zu verringern und übersichtlich zu halten, wurden nur Monografien berücksichtigt, die nach Autopsie katalogisiert worden sind. Die Erscheinungsjahre wurden ebenfalls von 2011 bis 2015 eingeschränkt.

Von der Analyse ausgeschlossen wurden Datensätze, die Originalschrift enthalten. Bei solchen Datensätzen besteht die Möglichkeit, dass für die Transliteration automatische Tools wie das Transliterationstool des WinIBWs verwendet worden sind. Die Verwendung solcher Tools könnte auf die Transliterationsergebnisse eine Auswirkung haben, die nicht zuverlässig genug eingeschätzt werden kann.

Ebenfalls nicht berücksichtigt wurden Titel, die nur aus einem Wort bzw. einem Namen bestehen und dabei keinen Titelzusatz aufweisen, da der Transliterationsbedarf hier sehr niedrig ist. Außerdem ausgeschlossen wurden Datensätze, die Russisch nur als eine von mehreren Sprachen aufweisen und deren Titel in einer anderen Sprache als Russisch vorliegt.

Im Ergebnis der beschriebenen Überlegungen wurde die folgende Suchanfrage formuliert:

“f spr rus and bbg aau and jhr XXXX not cod cyrl”.

An Stelle von XXXX wurde dabei der jeweilige zu analysierende Jahrgang eingefügt. Um die Trefferliste etwas überschaubarer zu halten, wurde die Stichprobe getrennt für jedes Erscheinungsjahr abgerufen und analysiert.

Während der Analyse wurde in erster Linie nach Transliterationsfehlern gesucht, und zwar nach der Verwendung von falschen Buchstaben (fehlende oder überschüssige Buchstaben, falsche Groß- und Kleinschreibung usw.). Die festgestellten Fehler wurden dabei nach Fehlerarten unterschieden.

In Bezug auf die Kategorisierung der Tippfehler orientierte sich die Analyse an der Klassifikation von Gentner et al. Diese Autoren widmeten ihren Beitrag detailliert dem Prozess des Tippens und definierten eine Reihe von Grundbegriffen, die mit dieser Tätigkeit verbunden sind und zu einer detaillierten Beschreibung und Untersuchung des Tippens verwendet werden könnten. Darüber hinaus schlugen sie eine Klassifikation der Tippfehler vor. Da die Autoren bei ihrer Klassifikation nicht nur Tippfehler an sich, sondern auch den Prozess des Tippens, also des Schreibens mit einer Schreibmaschine oder Tastatur, und die Besonderheiten der Tippfehler je nach der Position der Finger bei Anwendung eines Zehnfingersystems berücksichtigten, wurde ihre Klassifikation für die Zwecke der vorliegenden Arbeit angepasst.⁷¹

Somit wurden folgende Fehlerarten formuliert:

- Vertauschung (die Reihenfolge von zwei aufeinanderfolgenden Buchstaben wird vertauscht);
- Auslassung (ein oder mehrere Buchstaben werden ausgelassen);
- Weglassung der diakritischen Zeichen (diakritische Zeichen werden ausgelassen und es wird nur der Grundbuchstabe verwendet);
- Einfügen (ein oder mehrere falsche Buchstaben werden eingefügt);

⁷¹ Vgl. Gentner, Donald R. et al. (1983): A glossary of terms including a classification of typing errors, S. 42–43.

- Ersetzung (ein oder mehrere Buchstaben wird bzw. werden durch einen oder mehrere falsche Buchstaben ersetzt);
- Doppelung (ein Buchstabe oder eine Buchstabenkombination wird fälschlicherweise doppelt geschrieben);
- falsche Groß- und Kleinschreibung;
- Verwendung eines falschen Transliterationsstandards.

Während der Durchführung der Analyse wurden die festgestellten Tippfehler kategorisiert und die Fehlerart im Analyseprotokoll vermerkt. Das Analyseprotokoll diente somit zur Dokumentation der Datenerhebung und zur anschließenden Ableitung und Formulierung der Analyseergebnisse. Im Protokoll wurde jeder geprüfte Datensatz dokumentiert, und zwar unabhängig davon, ob er einen Tippfehler enthielt oder korrekt war. Für jeden Datensatz wurden folgende Informationen festgehalten:

- laufende Nummer des Datensatzes in der Stichprobe;
- PPN des Datensatzes;
- Tippfehlerstatus des Datensatzes;
- Art des Tippfehlers (im Falle eines Tippfehlers);
- sonstige Auffälligkeiten, falls vorhanden;
- fehlerhafte Schreibweise (im Falle eines Tippfehlers);
- korrekte Schreibweise (im Falle eines Tippfehlers);
- Schreibweise im Original⁷² sowie Kommentare, falls erforderlich,
- Bibliotheken, die ihren Lokal- bzw. Exemplardatensatz an den geprüften Datensatz angehängt und somit für die Erstellung des Datensatzes zuständig sind oder an seiner Erstellung eventuell beteiligt waren.

Für die Ausfüllung des Feldes „Art des Tippfehlers“ waren folgende Möglichkeiten vorgesehen:

1. Vertauschung;
2. Auslassung;
3. Weglassung der diakritischen Zeichen;
4. Einfügen;
5. Ersetzung;
6. Doppelung;
7. falsche Groß- und Kleinschreibung;
8. Verwendung eines falschen Transliterationsstandards.

Das Analyseprotokoll wird im Anhang A zu dieser Bachelor-Arbeit in elektronischer Form aufgeführt.

⁷² Die Schreibweise wurde vorzugsweise aus den OPACs der Russischen Nationalbibliothek und der Russischen Staatsbibliothek oder aus dem Buchhandel kopiert, und zwar in der Form, wie sie dort aufgeführt war.

3.1.3 Einschränkungen

Die angewendete Analysemethode verfügt über einige Einschränkungen.

Die Rekonstruktion der bibliografischen Angaben (Retransliteration) ins Russische erfolgte ausschließlich mental („im Kopf“) und unterliegt dabei einem Fehlriskio, das aufgrund der menschlichen Natur für alle rein intellektuelle Tätigkeiten charakteristisch ist. Dieses Risiko wird allerdings als akzeptabel eingestuft. Da die Verfasserin Russisch als Muttersprache beherrscht und über einen Universitätsabschluss in Linguistik sowie eine jahrelange Berufserfahrung als Übersetzerin verfügt, verfügt sie über genug Kompetenzen, um dieses Risiko gering zu halten. Um eine falsche Einstufung der Datensätze als fehlerhaft zu vermeiden, wurden im Analyseprotokoll sowohl die falsche als auch die richtige Schreibweise festgehalten. Zu diesem Zweck wurden die OPACs der Russischen Nationalbibliothek und der Russischen Staatsbibliothek⁷³ sowie andere mögliche Quellen zum Abgleich der bibliografischen Angaben hinzugezogen. Es wurden außerdem Datensätze aussortiert, wenn die entsprechenden bibliografischen Angaben nicht zuverlässig genug in die russische Sprache zurück transliteriert werden konnten.

Bei der Analyse wurde außerdem kein direkter Vergleich der Datensätze und der jeweiligen für die Formalerschließung verwendeten originalen Informationsquellen (Titelblatt u. a.) durchgeführt, da diese meistens nicht zur Verfügung standen. Somit beruht die Fehleranalyse auf der Annahme, dass die Originale der Monografien frei von Fehlern waren. Es wird angenommen, dass die Fehler bei der Dateneingabe während des Katalogisierungsvorgangs entstanden sein müssten. Da Verlage üblicherweise über ausreichende Qualitätssicherungsmaßnahmen verfügen, um Tipp- und Druckfehler insbesondere auf den Titelblättern zu vermeiden, wird dieses Risiko ebenfalls als geringfügig eingeschätzt.

Außerdem bestand wie bei jedem Forschungsvorhaben ein Risiko von unvorhersehbaren Ergebnissen. Es könnte vorkommen, dass die geplante Analyse keine Antworten zumindest auf einige der Forschungsfragen liefert. Die Forschungsfrage in Bezug auf die prozentuale Anzahl der fehlerhaften Datensätze scheint eher unproblematisch zu sein. Es könnte allerdings vorkommen, dass die Fehler sich schlecht kategorisieren lassen oder dass sich keine wiederholenden Muster identifizieren lassen, sodass keine sinnvollen Schlüsse gezogen werden können. Da negative Forschungsergebnisse ein Teil der Forschungsarbeit und für die Wissenschaft ebenfalls von Bedeutung sind, wird im Rahmen der vorliegenden Bachelor-Arbeit auch ein derartiges Ergebnis akzeptiert.

⁷³ Es handelt sich dabei um zwei (von insgesamt drei) Nationalbibliotheken der Russischen Föderation, die zu den größten Bibliotheken der Welt gehören. Die Russische Nationalbibliothek verfügt über 38,5 Millionen Medien, die Russische Staatsbibliothek über 48 Millionen Medien.

3.2 Ergebnisse der Analyse

3.2.1 Gesamtzahl der ermittelten und geprüften Datensätze

Im Ergebnis der im Punkt 3.2.1 formulierten Suchanfrage für die Jahrgänge 2011-2015 wurden insgesamt 27.902 Datensätze ermittelt. Da für die Analyse nur jeder fünfzigste Datensatz ausgewählt wurde, wurden insgesamt 552 Datensätze geprüft. Die Anzahl der durch die Suchanfrage jeweils für einen Jahrgang ermittelten und geprüften Datensätze ist in der nachfolgenden Tabelle 3 aufgeführt.

Tabelle 3. Anzahl der ermittelten und geprüften Datensätze nach Jahrgang

Jahrgang	Anzahl der ermittelten Datensätze (Trefferzahl)	Anzahl der geprüften Datensätze
2011	6.257	126
2012	5.948	119
2013	5.428	104
2014	5.445	107
2015	4.824	96
Insgesamt	27.902	552

Aus der Tabelle 3 lässt sich die Tendenz feststellen, dass die Anzahl der ermittelten und somit auch die der geprüften Datensätze höher ist, wenn der jeweilige Jahrgang weiter zurückliegt. Pro Jahrgang wurden insgesamt ca. 100 Datensätze geprüft. An dieser Stelle muss darauf hingewiesen werden, dass die Analyse im Januar 2020 durchgeführt wurde. Dementsprechend spiegeln die in der Tabelle 3 aufgeführten Zahlen den Stand von Januar 2020 wider. Es könnte vorkommen, dass die Zahl der Datensätze, die sich durch die formulierte Suchanfrage ermitteln lassen, sich zu einem späteren Zeitpunkt ändert, weil in der Verbunddatenbank neue Datensätze erstellt werden.

3.2.2 Gesamtzahl der fehlerhaften Datensätze und der festgestellten Fehler

Unter den 552 geprüften Datensätzen wurden insgesamt 80 Datensätze festgestellt, die Transliterationsfehler enthielten.⁷⁴ Die Gesamtzahl der festgestellten Transliterationsfehler belief sich dabei auf 110, da einige Datensätze mehrere Fehler auf einmal enthielten. Der Anteil der fehlerhaften Datensätze an der Gesamtzahl der geprüften Datensätze

⁷⁴ Wenn im Zuge dieses Kapitels von fehlerhaften Datensätzen bzw. Fehlern gesprochen wird, werden darunter Fehler verstanden, die zu einer in dieser Bachelor-Arbeit definierten Arten der Transliterationsfehler gehören.

liegt somit bei 14,5 %. Die Verteilung der fehlerhaften Datensätze auf die jeweiligen Jahrgänge wird in Tabelle 4 aufgeführt.

Tabelle 4. Verteilung der fehlerhaften Datensätze auf jeweilige Jahrgänge

Jahrgang	Anzahl der geprüften Datensätze	Anzahl der fehlerhaften Datensätze	Anteil der fehlerhaften Datensätze in Prozent
2011	126	21	16,7
2012	119	15	12,6
2013	104	8	7,7
2014	107	21	19,6
2015	96	15	15,6
Insgesamt	552	80	14,5

Der minimale Anteil der fehlerhaften Datensätze an der Gesamtzahl der geprüften Datensätze im jeweiligen Jahrgang liegt bei 7,7 %, der maximale Anteil bei 19,6 %. Aus der Verteilung der fehlerhaften Datensätze auf jeweilige Jahrgänge lassen sich keine allgemeinen Schlussfolgerungen ziehen. So kann zum Beispiel nicht behauptet werden, dass die „älteren“ Jahrgänge weniger Transliterationsfehler aufweisen, weil in der Zwischenzeit mehr Bibliotheken die jeweilige Monografie in ihren Bestand übernommen, ihre Exemplardaten an den entsprechenden Datensatz in der Verbunddatenbank angehängt und dabei eventuelle Fehler korrigiert haben. Die Verteilung der fehlerhaften Datensätze auf jeweilige Jahrgänge scheint eher einen zufälligen Charakter zu haben.

3.2.3 Aufschlüsselung der Datensätze nach besitzender Bibliothek

Als besitzende Bibliotheken wurden Bibliotheken erfasst, die in der Verbunddatenbank ihre Exemplardaten an den entsprechenden Datensatz angehängt haben und somit die jeweilige Monografie in ihrem Bestand haben.

Die Aufschlüsselung der Datensätze nach besitzender Bibliothek hat auf keinen Fall zum Ziel, die „beste“ und die „schlechteste“ Bibliothek in Bezug auf ihre Fehlerquote zu bestimmen. Sie dient dazu, eventuelle Muster und Abhängigkeiten festzustellen, zum Beispiel in Bezug darauf, wie oft in einer Bibliothek russischsprachige Monografien katalogisiert werden und wie viele Fehler dabei entstehen. Es könnte vermutet werden, dass große Bibliotheken, die einen größeren Bestand an russischsprachigen Neuerwerbungen und eine größere Katalogisierungsabteilung haben, eher die Möglichkeit haben, mit der Erschließung Fachkräfte zu beauftragen, die bestimmte Russischkenntnisse haben. Somit würde ihre Fehlerquote niedriger sein.

Leider lässt sich nicht feststellen, ob und wie oft Bibliotheken beim Anhängen ihres Exemplardatensatzes bereits vorhandene Transliterationsfehler entdecken und berichtigen. Wenn mehrere Bibliotheken ihre Exemplardaten an einen Datensatz angehängt haben, lässt sich dementsprechend auch nicht eindeutig feststellen, welche Bibliothek für die Korrektheit oder Fehlerhaftigkeit der Transliteration zuständig ist. Aus diesem Grund wurden für die Analyse in diesem Punkt Datensätze aussortiert, die Exemplardaten von mehreren Bibliotheken beinhalten. Bei den restlichen Datensätzen wird angenommen, dass eine Bibliothek den jeweiligen Titel katalogisiert hat, wenn sie als Einzige im Verbund ihn besitzt. Dabei wurde nicht berücksichtigt, ob diese Bibliothek das Katalogisat von null auf erstellt hat oder ob unter Umständen Fremddaten als Basis vorlagen.

Auf diese Weise wurden insgesamt 456 Datensätze ausgewählt, die Exemplardaten von einer einzigen Bibliothek aufweisen. Davon waren 383 korrekt, 73 enthielten Transliterationsfehler. Bibliotheken, die nur einen korrekten oder fehlerhaften Datensatz angelegt haben, wurden wegen unzureichender Datenmenge ausgeschlossen. Außerdem ausgeschlossen wurden die Goethe-Institute. Die Goethe-Institute werden in K10plus unter einer ILN (Internal Library Number) geführt. Tatsächlich handelt es sich aber um unterschiedliche Bibliotheken an unterschiedlichen Standorten, die eine eigene ELN (External Library Number) haben. Da bei der Datenerhebung nur die ILN berücksichtigt wurde und einzelne ELN nicht erfasst wurden, wäre für die Berücksichtigung der Goethe-Institute eine erneute Aufschlüsselung nach ELNs notwendig, auf die im Rahmen dieser Bachelor-Arbeit jedoch verzichtet wurde.

Unter Berücksichtigung der aufgeführten Überlegungen verringerte sich die Zahl der Datensätze auf 427. Die restlichen Daten wurden in Tabelle 5 nach der Gesamtzahl der durch eine Bibliothek angelegten Datensätze sowie der Anzahl an korrekten und fehlerhaften Aufnahmen aufgeschlüsselt. Auf Grundlage dieser Zahlen wurde die Fehlerquote berechnet.

Tabelle 5. Aufschlüsselung der Datensätze nach besitzenden Bibliotheken

Besitzende Bibliothek	Angelegte Datensätze insgesamt	davon korrekt	davon fehlerhaft	Fehlerquote in Prozent
Staatsbibliothek zu Berlin	188	163	25	13,2
ZBW Kiel	55	47	8	14,5
Niedersächsische Staats- und Universitätsbibliothek Göttingen	30	25	5	16,6
TIB Hannover	30	28	2	6,6
Stadtbibliothek Chemnitz	15	12	3	20

Besitzende Bibliothek	Angelegte Datensätze insgesamt	davon korrekt	davon fehlerhaft	Fehlerquote in Prozent
ULB Halle	11	9	2	18,1
SLUB Dresden	9	9	0	0
UB Freiburg	9	8	1	11,1
UB Heidelberg	8	7	1	12,5
UB Tübingen	8	5	3	37,5
Bibliotheken Saarland	7	7	0	0
UB Kiel	6	5	1	16,6
UB Konstanz	6	6	0	0
UB Leipzig	6	4	2	33,3
Forschungsstelle Osteuropa an der Universität Bremen	5	5	0	0
Leibniz Institut für Länderkunde	4	2	2	50
UB Greifswald	4	4	0	0
TU Bergakademie Freiberg	3	3	0	0
UB Mannheim	3	2	1	33,3
WLB Stuttgart	3	3	0	0
Staatliche Museen Berlin	2	1	1	50
Stadtbibliothek Braunschweig	2	2	0	0
Landesmuseum Württemberg	2	0	2	100
Datensätze insgesamt	416			

Tabelle 5 zeigt, dass fast die Hälfte (45 %) der Datensätze von der Staatsbibliothek zu Berlin stammt. Von insgesamt 416 Datensätze besitzt sie 188 russischsprachige Monografien als Einzige im Verbund. Die ZBW Kiel befindet sich auf Platz zwei und legte 55 Datensätze an. Danach folgen die Niedersächsische Staats- und Universitätsbibliothek Göttingen sowie die TIB Hannover mit 30 Datensätzen.

Je niedriger die Zahl der angelegten Datensätze, desto uneinheitlicher werden die Zahlen. Die Fehlerquote rangiert zwischen 0 % und 100 %, was in einigen Fällen unter anderem auf die niedrige Gesamtzahl der durch die jeweilige Bibliothek angelegten Datensätze zurückzuführen ist. Wenn eine Bibliothek nur zwei Datensätze angelegt hat und

beide korrekt oder fehlerhaft waren, können daraus keine belastbaren Schlussfolgerungen gezogen werden. Es könnte sein, dass beide Titel entweder sehr einfach oder sehr kompliziert zum Transliterieren waren, was kein verlässliches Gesamtbild ermöglicht. Für eine angemessene Analyse des in diesem Unterpunkt beleuchteten Aspekts ist eine viel größere Stichprobe erforderlich, was im Rahmen einer Bachelor-Arbeit aus zeitlichen Gründen nicht gewährleistet werden kann.

Wenn man allerdings nur die Statistik der ersten sechs Bibliotheken betrachtet, die alle mehr als zehn Datensätze angelegt haben, kann man sich zumindest einen ersten Eindruck bilden. Die TIB Hannover weist eine im Vergleich niedrigere Fehlerquote in Höhe von 6,6 % auf. Die Fehlerquoten der anderen fünf Bibliotheken sind vergleichbar und liegen zwischen 13,2 % bis 20 % mit einem Durchschnitt von 16,5 %. Diese Zahl hat eine ähnliche Größenordnung wie die Fehlerquote der gesamten Stichprobe in Höhe von 14,5 % (Anteil der fehlerhaften Datensätze an der Gesamtzahl der geprüften Datensätze, siehe Punkt 3.2.2), was für ihre tatsächliche Glaubwürdigkeit spricht.

Angesichts der beschränkten Stichprobe kann weder bestätigt noch widerlegt werden, dass ein größerer Bestand an russischsprachigen Neuerwerbungen mit einer niedrigeren Fehlerquote bei Transliteration korreliert. Es kann jedoch festgehalten werden, dass die Staatsbibliothek zu Berlin im SWB-GBV-Verbund deutlich mehr russischsprachige Monografien erfasst als andere Bibliotheken. Die Fehlerquoten der großen Bibliotheken, die eine höhere Anzahl an Datensätzen angelegt haben, sind allerdings vergleichbar. Unter den Bibliotheken, die in der Stichprobe weniger oft Datensätze angelegt haben, ist die Fehlerquote sehr unterschiedlich. Manche Bibliotheken weisen dabei eine höhere Fehlerquote auf. Da die absoluten Zahlen hier jedoch niedrig sind, lassen sich keine Schlüsse zu den möglichen Gründen ziehen.

3.2.4 Aufschlüsselung der festgestellten Fehler nach Fehlerart

Im Laufe der Datenanalyse wurden die festgestellten Fehler nach Fehlerarten kategorisiert. Die nachfolgende Tabelle 6 zeigt, wie oft die einzelnen Fehlerarten in der Stichprobe vorkommen. Wenn in einem Datensatz mehrere Fehler derselben Fehlerart festgestellt wurden, wurden diese Fehler jeweils getrennt gezählt. Anschließend werden die Zahlen nach unterschiedlichen Fehlerarten in der Abbildung 1 als Diagramm veranschaulicht.

Tabelle 6. Aufschlüsselung der festgestellten Fehler nach Fehlerart

Fehlerart	Anzahl der Fehler in der Stichprobe	Anteil der Fehler an der Gesamtfehlerzahl in Prozent
Ersetzung	30	27,3
Falsche Groß- und Kleinschreibung	28	25,5

Fehlerart	Anzahl der Fehler in der Stichprobe	Anteil der Fehler an der Gesamtfehlerzahl in Prozent
Auslassung	17	15,5
Weglassung der diakritischen Zeichen	15	13,6
Verwendung eines falschen Transliterationsstandards	8	7,3
Einfügen	6	5,5
Falsche diakritische Zeichen	3	2,7
Doppelung	2	1,8
Vertauschung	1	0,9
Insgesamt	110	

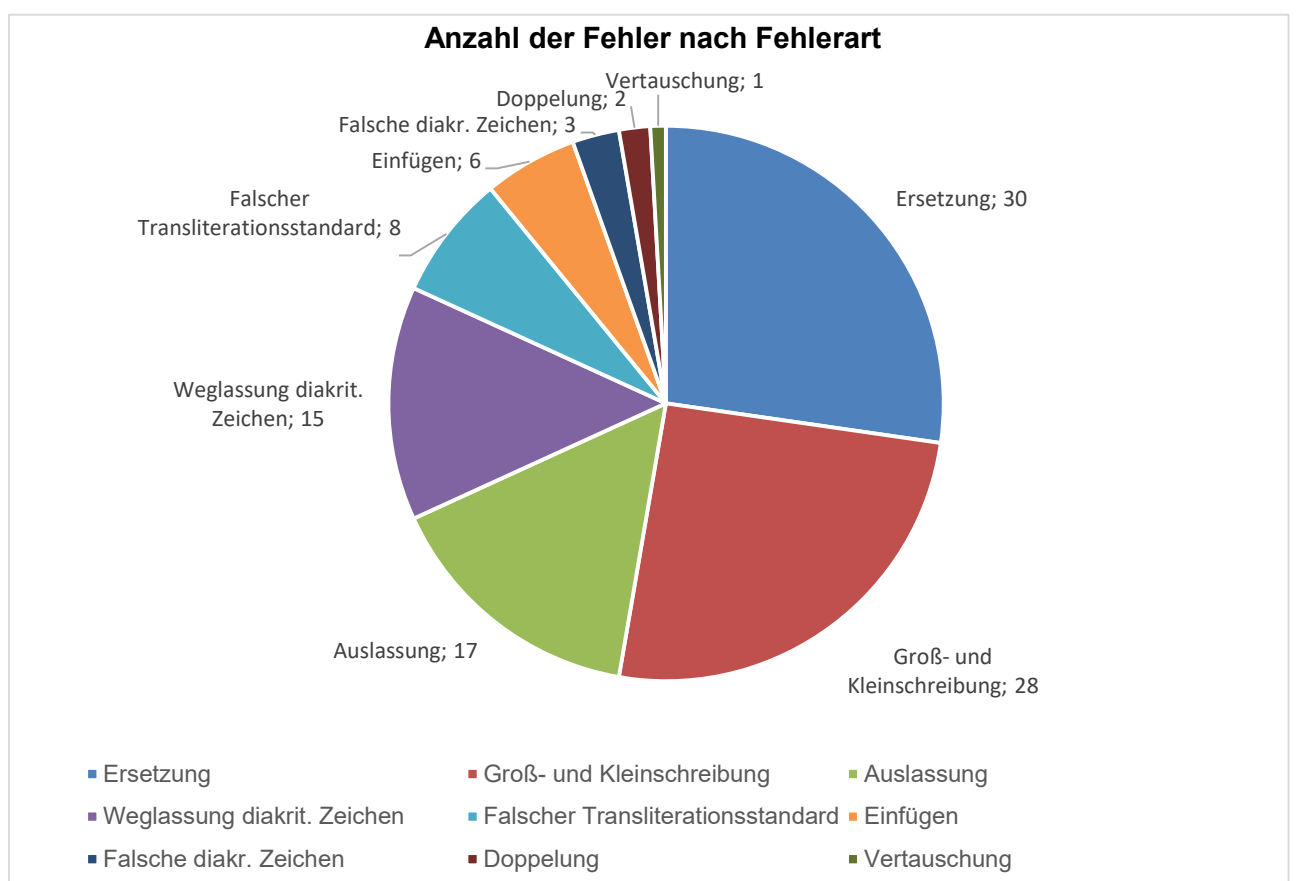


Abbildung 1. Anteile der einzelnen Fehlerarten an der Gesamtfehlerzahl

Somit lässt sich feststellen, dass die Anzahl der Fehler je nach Fehlerart sehr unterschiedlich ist. Am häufigsten kommen Ersetzung und falsche Groß- und Kleinschreibung vor, Verwendung falscher diakritischer Zeichen, Vertauschung und Doppelung sind jedoch selten. Die anderen Fehlerarten befinden sich im mittleren Bereich. In Tabelle 6 wird die Fehlerart „falsche diakritische Zeichen“ aufgeführt, die in der ursprünglichen Formulierung der möglichen Fehlerarten im Punkt 3.1.2. nicht erwähnt wurde. Im Zuge der Datenanalyse hat sich jedoch die Notwendigkeit für die Einführung dieser zusätzlichen Fehlerart erwiesen. Diese und andere in Tabelle 6 aufgeführten Fehlerarten werden in den folgenden Unterpunkten einzeln analysiert und ausführlicher erläutert, und zwar in der Reihenfolge, wie die einzelnen Fehlerarten im Punkt 3.1.2 sowie im Analyseprotokoll nummeriert wurden.

3.2.4.1 Vertauschung

Bei der Planung der Datenanalyse wurde im Punkt 3.1.2. als Vertauschung die Art von Transliterationsfehlern definiert, bei der die Reihenfolge von zwei aufeinanderfolgenden Buchstaben vertauscht wird.

Von insgesamt 110 festgestellten Fehlern handelte es sich lediglich in einem einzigen Fall um eine Vertauschung. Somit beträgt der Anteil dieser Fehlerart an der Gesamtzahl der festgestellten Fehler 0,9 % und ist minimal. Der Anteil der Vertauschung an der Gesamtzahl der fehlerhaften Datensätze liegt bei 1,25 % Prozent.

In der nachfolgenden Tabelle 7 wird die einzige festgestellte Vertauschung aufgeführt:

Tabelle 7. Vertauschung in der Stichprobe

PPN	Fehlerhafte Schreibweise	Korrekte Schreibweise	Schreibweise in Originalschrift
805407170	Četyre tuberozy\$ d poézija, proza i d armaturgija zabytych avtorov Serebrjanogo veka	Četyre tuberozy\$ d poézija, proza i d ramaturgija zabytych avtorov Serebrjanogo veka	Четыре туберозы\$ d позэзия, проза и д рама ^т тургия забытых авторов Серебряного века

Der Tabelle 7 kann entnommen werden, dass im jeweiligen Datensatz alle Buchstaben richtig erkannt und transliteriert worden sind, jedoch die Reihenfolge der Buchstaben „r“ und „a“ im Wort „dramaturgija“ vertauscht wurde. Es könnte angenommen werden, dass es sich dabei eher um einen Flüchtigkeitsfehler handelt, der unbemerkt geblieben ist. Die Tatsache, dass der Fehler dieser Art in der Stichprobe nur ein einziges Mal vorkommt, spricht dafür, dass Vertauschung eine unbedeutende Rolle in der Transliteration spielt. Diese Fehlerart ist jedoch retrievalrelevant.

3.2.4.2 Auslassung

Als Auslassung wird die Fehlerart verstanden, wenn bei der Transliteration ein oder mehrere Buchstaben fälschlicherweise ausgelassen werden.

In der Stichprobe von insgesamt 552 geprüften Datensätzen wurden 16 Datensätze festgestellt, die eine Auslassung enthalten. Die tatsächliche Anzahl der Auslassungen liegt dabei bei 17. Der Anteil der Datensätze, die eine Auslassung enthalten, an der Gesamtzahl der fehlerhaften Datensätze beträgt 20 %, der Anteil der Auslassung an der Gesamtzahl der festgestellten Fehler dementsprechend 15,5 %.

Einige Beispiele der einzelnen Datensätze, die eine Auslassung enthalten, werden zur Veranschaulichung in Tabelle 8 aufgeführt.

Tabelle 8. Beispiele von Auslassungen

PPN	Fehlerhafte Schreibweise	Korrekte Schreibweise	Schreibweise in Originalschrift
1619518910	Jugo-zapadne predmest'ja Peterburga\$ d istorija, arhitektura, dač n ja žizn'	Jugo-zapadne predmest'ja Peterburga\$ d istorija, arhitektura, dač n a žizn'	Юго-западные предместья Петербурга\$ d история, архитектура, дач n ая жизнь
846148714	Glagol'naja leksika russkich govorov Volgo- Kam j a\$ d semantika i slovoproizvodstvo	Glagol'naja leksika russkich govorov Volgo- Kam j 'a\$ d semantika i slovoproizvodstvo	Глагольная лексика русских говоров Волго- Кам ь \$ d семантика и словопроизводство
827944357	Rodonačal'niki rossijskogo pred r inimatel'stva	Rodonačal'niki rossijskogo pred p rinimatel'stva	Родоначалъники российского пред p ринимательства
770130933	Pamjat t serdca\$ d stichi, pesni, "Ispoved" žizni, stat'i	Pamjat' t serdca\$ d stichi, pesni, "Ispoved" žizni, stat'i	Памят ь сердца\$ d стихи, песни, "Исповедь" жизни, статьи
663662559	Lečenje i obučenie sotrudnikov\$ d ak dokumental'no oformit', učest t , oplatit' i kompensirovat' s minimal'nymi	Lečenje i obučenie sotrudnikov\$ d ak dokumental'no oformit', učest t , oplatit' i kompensirovat' s minimal'nymi	Лечение и обучение сотрудников\$ d ак документально оформить, учест ь , оплатить и компенсировать с минимальными

PPN	Fehlerhafte Schreibweise	Korrekte Schreibweise	Schreibweise in Originalschrift
	nalogovymi riskami ; razrabotka temy, kommentarii i raz"jasnenija ...	nalogovymi riskami ; razrabotka temy, kommentarii i raz"jasnenija ...	налоговыми рисками ; разработка темы, комментарии и разъяснения...

Die festgestellten Auslassungen wurden nach ausgelassenen Buchstaben aufgeschlüsselt. Die in Tabelle 9 unten aufgeführte Analyse der Datensätze in Hinblick darauf, welche Buchstaben wie oft ausgelassen wurden, liefert einige interessante Erkenntnisse.

Tabelle 9. Ausgelassene Buchstaben und Häufigkeit ihrer Auslassung

Russische(r) Buchstabe(n)	Lateinische(r) Buchstabe(n)	Anzahl der Auslassungen
ь	,	10
й	j	2
во	vo	2
п	p	1
л	l	1
а	a	1
Insgesamt		17

Insgesamt wurden sechs unterschiedliche Buchstaben oder Buchstabenkombinationen ausgelassen. Fast zwei Drittel (10 von 17) der Auslassungen entfallen dabei auf den russischen Buchstaben „ь“. Am zweithäufigsten wurde zwei Mal der russische Buchstabe „й“ ausgelassen, mit der gleichen Häufigkeit tritt die Auslassung einer Kombination von zwei Buchstaben „во“ auf (russische Buchstaben „в“ und „о“). Drei weitere Buchstaben „а“, „п“ und „л“ wurden nur jeweils ein Mal ausgelassen.

Der Buchstabe „ь“ wird im Russischen als „weiches Zeichen“ bezeichnet. Er selbst wird nicht ausgesprochen und ist nur ein Zeichen davon, dass der jeweils vorangehende Buchstabe weich ausgesprochen werden soll. Von der äußerlichen Schreibweise ähnelt er sich dabei den Buchstaben „ъ“, „Ѣ“ und „Ѥ“. Hier könnte von einer gewissen Verwechslungsgefahr ausgegangen werden, besonders zwischen den Buchstaben „ь“ und „ъ“, die für Personen ohne Sprachkenntnisse sehr ähnlich aussehen. In den 10 analysierten Datensätzen kam es jedoch nicht zu einer Verwechslung, sondern zu einer Auslassung des Buchstabens „ь“. Wie die eigenen Erfahrungen der Verfasserin zeigen, wäre gerade eine Person, die Russisch spricht, eher dazu geneigt, den Buchstaben „ь“ auszulassen als

zu vertauschen. Da er nicht ausgesprochen wird, geht er bei der Transliteration auch leichter „verloren“. Die Auslassung dieses Buchstabens anstatt seiner Verwechslung könnte dafür sprechen, dass die entsprechenden Datensätze von Personen mit Russischkenntnissen aufgenommen wurden. Um diese Vermutung zu bestätigen, wurden die Datensätze nach der Bibliothek analysiert, die den Datensatz erfasst hat. Die Ergebnisse dieser Analyse werden in der Tabelle 10 unten zusammengefasst.

Tabelle 10. Auslassung des Buchstabens „b“ nach erfassender Bibliothek

Bibliothek	Anzahl der Auslassungen des Buchstabens „b“
Staatsbibliothek zu Berlin	6
Goethe-Institute	3
Niedersächsische Staats- und Universitätsbibliothek Göttingen	1
Insgesamt	10

Von insgesamt 10 Auslassungen befinden sich 6 in Datensätzen, die von der Staatsbibliothek zu Berlin erfasst worden sind. Zwei Datensätze, die jedoch insgesamt drei Auslassungen enthalten, stammen von den Goethe-Instituten. Nur eine Auslassung stammt aus dem Datensatz der Niedersächsischen Staats- und Universitätsbibliothek Göttingen.

Der eine Datensatz der Goethe-Institute (PPN 1614584192) mit zwei Auslassungen wurde dabei vom Goethe-Institut Minsk als einziger besitzender Bibliothek erfasst. Bei dem zweiten Datensatz (PPN 1602551049) besitzen mehrere Goethe-Institute ein Exemplar der Monografie, die früheste Eintragung stammt jedoch ebenfalls vom Institut in Minsk. In Weißrussland ist Russisch neben Weißrussisch eine der beiden Staatssprachen und sogar die am meisten benutzte Verkehrssprache (circa 85 %)⁷⁵. Daher ist die Wahrscheinlichkeit, dass das katalogisierende Personal Russisch spricht, relativ hoch. In Bezug auf die Staatsbibliothek zu Berlin kann aufgrund ihrer eher niedrigen Fehlerquote und einer sehr hohen Anzahl an Katalogisaten russischsprachiger Monografien (siehe Punkt 3.2.3) ebenfalls vermutet werden, dass russischsprachige Monografien in der Bibliothek vom Personal katalogisiert werden, welches über Russischkenntnisse verfügt.

Die Annahme, dass der Buchstabe „b“ besonders leicht von Personen mit Russischkenntnissen ausgelassen wird, lässt sich indirekt durch einige der Datensätze bestätigen. Der von der Staatsbibliothek zu Berlin stammende Datensatz PPN 663662559 (siehe Tabelle 8) weist einen Haupttitel mit einem sehr langen Titelzusatz auf, bestehend

⁷⁵ Vgl. Auswärtiges Amt (2019): Belarus: Überblick.

aus langen Wörtern mit vielen „Transliterationstücken“ wie diakritischen Zeichen, mehrfach vorkommendem Buchstaben „б“ und sogar dem Buchstaben „ѣ“. Die ansonsten fehlerfreie Transliteration so eines anspruchsvollen Titels und lediglich eine Auslassung des insgesamt sechs Mal vorkommenden Buchstabens „б“ wäre vermutlich insbesondere für eine Person mit entsprechenden Sprachkenntnissen charakteristisch.

Auslassung gehört neben der Vertauschung zu den Fehlerarten, die einen negativen Einfluss auf das Retrieval haben.

3.2.4.3 Weglassung der diakritischen Zeichen

Unter der Weglassung der diakritischen Zeichen werden Fehler verstanden, wenn diakritische Zeichen ausgelassen werden und nur der jeweilige Grundbuchstabe verwendet wird.

In der Stichprobe wurden 14 Datensätze festgestellt, die insgesamt 15 Fehler dieser Art aufweisen. Somit beträgt der Anteil der Datensätze, die eine Weglassung der diakritischen Zeichen enthalten, an der Gesamtzahl der fehlerhaften Datensätze 17,5 %, der Anteil der entsprechenden Fehler an der Gesamtfehlerzahl liegt bei 13,6 %.

In der nachfolgenden Tabelle 11 werden einige Beispiele der einzelnen Datensätze aufgeführt, die eine Weglassung diakritischer Zeichen enthalten.

Tabelle 11. Beispiele der Weglassung diakritischer Zeichen

PPN	Fehlerhafte Schreibweise	Korrekte Schreibweise	Schreibweise in Originalschrift
834209853	Proischoždenie sovremennogo narodno-pravovogo gosudarstva \$distoričeskij očerk konstitucionnych ucreždenij i učenij do serediny XIX veka	Proischoždenie sovremennogo narodno-pravovogo gosudarstva \$distoričeskij očerk konstitucionnych učreždenij i učenij do serediny XIX veka	Происхождение современного народно-правового государства \$дисторический очерк конституционных учреждений и учений до середины XIX века
872150267	Kratkij spravocnik gal'vanotechnika	Kratkij spravočnik gal'vanotechnika	Краткий справочник гальванотехника
718796942	Baškirskaia poezija XIX - načala XX vekov\$dvoprosy žanrovych i idejno-	Baškirskaia poëzija XIX - načala XX vekov\$dvoprosy žanrovych i idejno-	Башкирская поэзия XIX - начала XX веков\$двопросы жанровых и идейно-

	chudožestvennych osobennosti <i>i</i>	chudožestvennych osobennosti <i>j</i>	художественных особенностей <i>й</i>
--	--	--	---

Die DIN 1460 sieht insgesamt sechs Buchstaben bzw. Buchstabenkombinationen vor, in denen diakritische Zeichen vorkommen, und zwar „ë“ für das russische „ё“, „ž“ für „ж“, „č“ für „ч“, „š“ für „ш“, „šč“ für „щ“ und „é“ für „э“. Die nachfolgende Tabelle 12 enthält die Aufschlüsselung, welche Buchstaben von der Weglassung der diakritischen Zeichen in der Stichprobe betroffen sind.

Tabelle 12. Aufschlüsselung der Weglassung diakritischer Zeichen

Buchstabe	Anzahl der Weglassungen diakritischer Zeichen
ë	8
č	4
ž	3
Insgesamt	15

Die aufgeführte Aufschlüsselung zeigt, dass diakritische Zeichen am häufigsten beim Buchstaben „ë“ ausgelassen werden, und zwar in mehr als 50 % der in der Stichprobe festgestellten Fehler dieser Art. Die andere Hälfte der Fehler wird fast zu gleichen Anteilen zwischen den Buchstaben „č“ und „ž“ verteilt.

Die Erkenntnis, dass diakritische Zeichen am häufigsten im Buchstaben „ë“ weggelassen werden, könnte an der gewissen Ähnlichkeit zu seinem Grundbuchstaben „e“ liegen. Das russische „e“ ist nämlich ein weich anzeigendes Gegenstück des hart anzeigenden Buchstaben „э“, dem Klang vom Letzten wird dabei ein „j“ vorne hinzugefügt (ausgesprochen wie „je“). Außerdem wird das russische „e“ in manchen Wörtern sogar wie ein „э“ ausgesprochen (wie z. B. „Интернет“ (Internet)). Dadurch könnte es leichter zu Situationen kommen, wenn das Fehlen der diakritischen Zeichen beim Buchstaben „ë“ während der Überprüfung des jeweiligen Datensatzes vor seinem Speichern nicht bemerkt wird.

Auffällig ist in der Stichprobe der Datensatz PPN 1645265862 mit dem Titel „S Dnem roždenija, Karlchen!“ („С днём рождения, Карлхен!“, übersetzt „Alles Gute zum Geburtstag, Karlchen!“). Hier wurde höchstwahrscheinlich das Tremazeichen (der horizontale Doppelpunkt) beim Buchstaben „ë“ ausgelassen. Im Online-Buchhandel sind nur das Buchcover sowie Bilder von den einzelnen Seiten zu finden, allerdings kein Titelblatt. Sowohl auf dem Cover des Buches (siehe Abbildung 2 unten) sowie im Text wird der Buchstabe „ë“ immer mit dem Trema geschrieben. Das Tremazeichen beim Buchstaben

„ë“ ist im Russischen nämlich fakultativ und nur in einer sehr begrenzten Reihe von Fällen wirklich erforderlich. Eins davon sind Texte für Kinder, die das Lesen erst erlernen. Daher ist davon auszugehen, dass das Trema beim Buchstaben „ë“ in diesem Kinderbuch durchgehend geschrieben wurde, auch auf dem Titelblatt. Da das Titelblatt leider nicht zur Verfügung steht, bleibt es allerdings eine Vermutung. Daher wurde dieser Sachverhalt nicht als ein Fehler berücksichtigt.



Abbildung 2. Buchcover zum Datensatz PPN 1645265862⁷⁶

Die Tatsache, dass in der Stichprobe keine weiteren Weglassungen des Tremas beim Buchstaben „ë“ festgestellt wurden, ist an sich nicht überraschend. Wie bereits erwähnt, wird dieses diakritische Zeichen im Russischen in den meisten Fällen nur fakultativ geschrieben und in der Praxis wird auf das Trema meistens verzichtet, wodurch der Buchstabe auch nur sehr selten in bibliografischen Angaben vorkommt.

In Bezug auf die Weglassung der diakritischen Zeichen ist zudem zu erwähnen, dass sie in der Praxis nicht retrievalrelevant sind. Da diakritische Zeichen bei Indexierung weggelassen werden und nur der Grundbuchstabe indexiert wird, hat diese Fehlerart keinen praktischen Einfluss auf das Retrieval.

3.2.4.4 Einfügen

Das Einfügen liegt vor, wenn in ein Wort ein falscher Buchstabe bzw. eine Buchstabenkombination eingefügt wird, der bzw. die an der Stelle nicht vorkommt. Diese Fehlerart

⁷⁶ Quelle der Abbildung: <https://www.labyrinth.ru/books/453161/> (letzter Zugriff: 05.02.2020)

sollte von Doppelung unterschieden werden, bei der ein Buchstabe oder eine Buchstabenkombination fälschlicherweise doppelt geschrieben wird (siehe Punkt 3.2.4.6).

Während der Datenanalyse wurden 6 Datensätze erfasst, die jeweils einen Fehler dieser Art enthielten. Die Zahl der fehlerhaften Datensätze und die Zahl der einzelnen Fehler waren in diesem Fall gleich. Somit liegt der Anteil der Datensätze, die ein Einfügen enthalten, an der Gesamtzahl der fehlerhaften Datensätze bei 7,5 %. Der Anteil der entsprechenden Fehler an der Gesamtfehlerzahl beträgt 5,5 %.

In der Stichprobe kam es zum Einfügen der Buchstaben „i“, „t“, „j“ sowie eines Bindestriches. Einige Beispiele sind in Tabelle 13 unten aufgeführt. Dabei kamen das Einfügen eines „i“ und eines „j“ zwei Mal vor, alle anderen Fälle waren Einzelfehler.

Tabelle 13. Beispiele für das Einfügen

PPN	Fehlerhafte Schreibweise	Korrekte Schreibweise	Schreibweise in Originalschrift
1546187553	Planeta ide <i>ij</i> \$dtransgraničnyj transfer znanij v oblasti privlečenija investicij dlja prigraničnogo turizma Pol'sha-Belarus'- Ukraina ot zamysla do realizacii	Planeta ide <i>j</i> \$dtransgraničnyj transfer znanij v oblasti privlečenija investicij dlja prigraničnogo turizma Pol'sha- Belarus'-Ukraina ot zamysla do realizacii	Планета иде <i>ей</i> \$dтрансграничный трансфер знаний в области привлечения инвестиций для приграничного туризма Польша – Беларусь – Украина от замысла до реализации
1478961678	Sofii - s ljubov'ju\$dk 80-letiju Sofii Asgatovny Gubajdulinoj; materialy meždunarodnoj naučno-praktičeskoj konferenc <i>tcii</i>	Sofii - s ljubov'ju\$dk 80-letiju Sofii Asgatovny Gubajdulinoj ; materialy meždunarodnoj naučno-praktičeskoj konferenc <i>ncii</i>	Софии - с любовью\$dk 80- летию Софии Асгатовны Губайдулиной ; материалы международной научно- практической конферен <i>нции</i>
661934357	Obščestvennye dviženij <i>a</i> na territorii Belarusi (XIX - načalo XX v.)\$distoriografija,	Obščestvennye dviženij <i>a</i> na territorii Belarusi (XIX - načalo XX v.)\$distoriografija,	Общественные движен <i>ия</i> на территории Беларуси (XIX - начало XX

	metodologičeskie principy issledovanija	metodologičeskie principy issledovanija	в.)\$dисториография, методологические принципы исследования
--	--	--	--

In Bezug auf das Einfügen des im Original nicht vorhandenen Bindestriches (letztes Beispiel in der Tabelle 13 oben) lässt sich festgehalten, dass die Position im Wort, wo er fälschlicherweise eingefügt worden ist, nicht den Anschein eines reinen Zufalls erweckt. Der Bindestrich wurde im Datensatz PPN 795228880 mit dem Titel „Sovremennaja rol' **e**konomiki Sibiri v narodno**o-c**hozjajstvennom komplekse Rossii“ („Современная роль экономики Сибири в народнохозяйственном комплексе России“, übersetzt „Die moderne Rolle der Wirtschaft von Sibirien für die russische Volkswirtschaft“) eingefügt. Auf dem Titelblatt ist dabei kein Bindestrich zu sehen (siehe Abbildung 3 unten). Außerdem wurde im Datensatz das diakritische Zeichen im Buchstaben „ё“ ausgelassen. Die korrekte Schreibweise wäre „Sovremennaja rol' **ë**konomiki Sibiri v narodno**o**chozjajstvennom komplekse Rossii“.

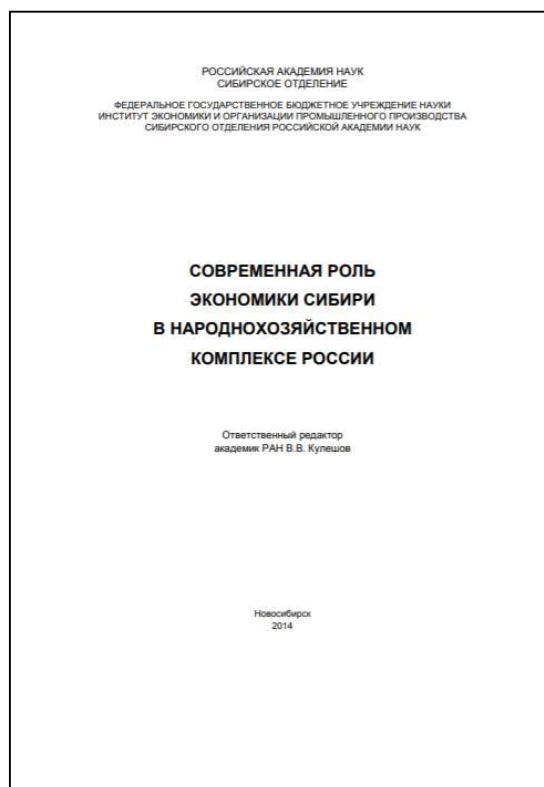


Abbildung 3. Titelblatt zum Datensatz PPN 795228880⁷⁷

⁷⁷ Quelle der Abbildung:
http://lib.ieie.su/docs/2014/SovremRolEconSib/Sovremennaya_Rol_Ekonomiki_Sibiri.pdf
 (letzter Zugriff: 19.02.2020)

Wie bereits erwähnt, ist die Position des Bindestriches nicht zufällig. Er wurde zwischen zwei sinngemäßen Komponenten eines zusammengesetzten Adjektivs eingefügt. Ein Bindestrich kann in der russischen Sprache in einer ähnlichen Position durchaus vorkommen, allerdings geben weder das Titelblatt noch die Rechtschreibregeln Anlass zu seiner Verwendung im betroffenen Wort des Haupttitels. Da die Platzierung des Bindestriches an dieser Stelle sinngemäß nachvollziehbar ist, könnte vermutet werden, dass der Datensatz von einer Person erstellt wurde, die Russisch gut beherrscht.

Da die Gesamtzahl des Einfügens und die Häufigkeit der einzelnen fälschlicherweise eingefügten Buchstaben niedrig sind, können in Bezug auf diese Fehlerkategorie keine Muster festgestellt werden. Generell lässt sich festhalten, dass diese Fehlerart einen negativen Einfluss auf das Retrieval hat.

3.2.4.5 Ersetzung

Als Ersetzung wurden Fehler definiert, wenn ein Buchstabe oder eine Buchstabenkombination durch einen falschen Buchstaben oder eine falsche Buchstabenkombination ersetzt wird.

In der Stichprobe wurden 26 Datensätze festgestellt, die insgesamt 30 Fehler dieser Art enthielten. Somit liegt der Anteil der Datensätze, die eine Ersetzung enthalten, an der Gesamtzahl der fehlerhaften Datensätze bei 32,5 %. Der Anteil der Ersetzungen an der Gesamtfehlerzahl der Stichprobe liegt bei 27,3 %. Ausgewählte Beispiele von festgestellten Ersetzungen sind in der Tabelle 14 aufgeführt.

Tabelle 14. Beispiele für Ersetzungen

PPN	Fehlerhafte Schreibweise	Korrekte Schreibweise	Schreibweise in Originalschrift
1687354200	Al'bom populjarn ich p'es \$d lja fortepiano	Al'bom populjarn ych p'es \$d lja fortepiano	Альбом популярных пьес \$d для фортепиано
670186481	Zapachi doždei \$d novelly i rasskazy	Zapachi doždej \$d novelly i rasskazy	Запахи дождей \$d новеллы и рассказы
160458954X	Ljubov' i logika prostran ctva \$d skul'ptura, ž ivopis'	Ljubov' i logika prostran stva \$d skul'ptura, ž ivopis'	Любовь и логика пространства \$d скульптура и живопись

Die festgestellten Fehler wurden nach Buchstaben analysiert, die fälschlicherweise ersetzt worden sind. Dabei wurde festgehalten, welcher Buchstabe gegen welchen anderen Buchstaben ersetzt worden ist. Die Aufschlüsselung der Ergebnisse wird in Tabelle 15 unten vorgestellt.

Tabelle 15. Aufschlüsselung der Ersetzungen

Ersetzter Buchstabe	Anzahl der Fehler
i anstatt j	5
c anstatt s	3
b anstatt d	2
i anstatt y	2
j anstatt i	2
y anstatt i	1
č anstatt s	1
` anstatt b	1
ň anstatt n`	1
ë anstatt e	1
e anstatt i	1
u anstatt i	1
r anstatt j	1
l anstatt j	1
f anstatt v	1
o anstatt y	1
f anstatt j	1
j anstatt `	1
b anstatt v	1
s anstatt z	1
i anstatt `	1
Insgesamt	30

Insgesamt kamen in der Stichprobe 21 Kombinationen vor, in denen ein Buchstabe durch einen anderen ersetzt worden ist. Es gab 16 Ersetzungen von Buchstaben, die

nur einmal vorgekommen sind. Lediglich fünf Buchstaben kamen mehrmals vor. Am häufigsten (fünf Mal) wurde ein „i“ anstatt „j“ geschrieben. Insgesamt befinden sich die Zahlen in Tabelle 15 in einem niedrigen Bereich, sodass sich keine eindeutigen Muster für die Ersetzung der einzelnen Buchstaben feststellen lassen.

Auffällig ist allerdings, dass der Buchstabe „i“ (russisches „и“) in der Tabelle 15 öfter vorkommt und insgesamt bei 13 Ersetzungen zu sehen ist. Fünf Mal wurde ein „i“ anstatt einem „j“ (russisches „й“), zwei Mal anstatt einem „y“ („ы“) und ein Mal anstatt einem „“ („ѣ“) geschrieben. Umgekehrt wurde ein „i“ zwei Mal durch ein „j“ und jeweils ein Mal durch ein „y“ (russisches „ы“), ein „e“ („е“) und ein „u“ („у“) fälschlicherweise ersetzt. Zwischen dem russischen „и“ und „й“ besteht dabei eine große grafische Ähnlichkeit, da die beiden Buchstaben sich nur durch ein diakritisches Zeichen unterscheiden. Es wäre möglich, dass sie wegen dieser Ähnlichkeit öfter fälschlicherweise gegeneinander ersetzt werden. Zwischen dem russischen „и“ und „ы“ besteht eine Ähnlichkeit in der Aussprache, denn „и“ ist eine weiche Variante von „ы“. Es könnte sein, dass die gegenseitige Ersetzung der beiden Buchstaben bei der Prüfung des Datensatzes vor seinem Abspeichern übersehen wird, vor allem von den Personen mit Russischkenntnissen, die die Wörter sich gedanklich „vorlesen“.

Die Ersetzung eines „“ (russisches „ѣ“) durch ein „i“ kommt in einem Vornamen vor, und zwar im Datensatz PPN 835081508 mit dem Titel „Sofija L'vovna Perovskaja\$ddžizň, ličnost', sud'ba“ („Софья Львовна Перовская\$дд жизнь, личность, судьба“, übersetzt „Sof'ja L'vovna Perovskaja\$дд Leben, Persönlichkeit, Schicksal“). Hier handelt es sich um einen Personennamen einer bekannten historischen Persönlichkeit, die allerdings „Sof'ja“ mit Vornamen hieß. Sofija und Sof'ja sind dabei Varianten desselben Vornamens, die Variante Sofija ist sogar gängiger. Dies könnte den Ersetzungsfehler erklären. Außerdem wurde im betroffenen Datensatz der gleiche Buchstabe „“ („ѣ“), allerdings diesmal in Kombination mit einem „n“ davor, durch ein „ñ“ ersetzt. Die korrekte Schreibweise wäre „Sof'ja L'vovna Perovskaja\$ддžizn', ličnost', sud'ba“.

Die restlichen Ersetzungen eines „i“ durch „e“ und „u“ erwecken den Eindruck, zufällig zu sein.

Bei Ersetzungen handelt es sich um eine Fehlerart, die retrievalrelevant ist.

3.2.4.6 Doppelung

Unter Doppelung werden im Rahmen dieser Bachelor-Arbeit Fehler definiert, wenn ein Buchstabe oder Buchstabenkombination fälschlicherweise doppelt geschrieben wird.

Im Rahmen der Datenanalyse wurden lediglich zwei Datensätze festgestellt, die jeweils einen Fehler dieser Art enthielten. Somit liegt der Anteil der Datensätze, die eine Doppelung enthalten, an der Gesamtzahl der fehlerhaften Datensätze lediglich bei 2,5 %. Der Anteil der einzelnen Fehler an der Gesamtzahl der Fehler in der Stichprobe liegt bei 1,8 % und ist somit eher unbedeutend.

Die erste der beiden festgestellten Doppelungen kommt im Datensatz PPN 161883679X mit dem Titel „Pianistka\$**dP**jessy“ („Пианистка\$**d**пьесы“, übersetzt „Die Klavierspielerin\$**d**Theaterstücke“) vor. Im Titelzusatz wurde der Buchstabe „s“ fälschlicherweise doppelt geschrieben. Die richtige Schreibweise wäre „Pianistka\$**dP**'esy“. In diesem aus nur einem Wort bestehenden Titelzusatz ist jedoch vieles missglückt, denn er enthält insgesamt drei Fehler. Neben der Doppelung liegen hier eine falsche Groß- und Kleinschreibung sowie eine Ersetzung (ein „j“ anstatt einem „‘“) vor.

Darüber hinaus ist es nicht nachvollziehbar, woher dieser Titelzusatz überhaupt stammt. Ein derartiger Titelzusatz ist weder im OPAC der Russischen Nationalbibliothek noch im OPAC der Russischen Staatsbibliothek zu finden. Die im Online-Buchhandel verfügbaren Bilder vom Buchcover (siehe Abbildung 4 unten), Titelblatt (Abbildung 5) sowie der Rückseite des Titelblattes (Abbildung 6) lassen ebenfalls keinen Titelzusatz erkennen⁷⁸.

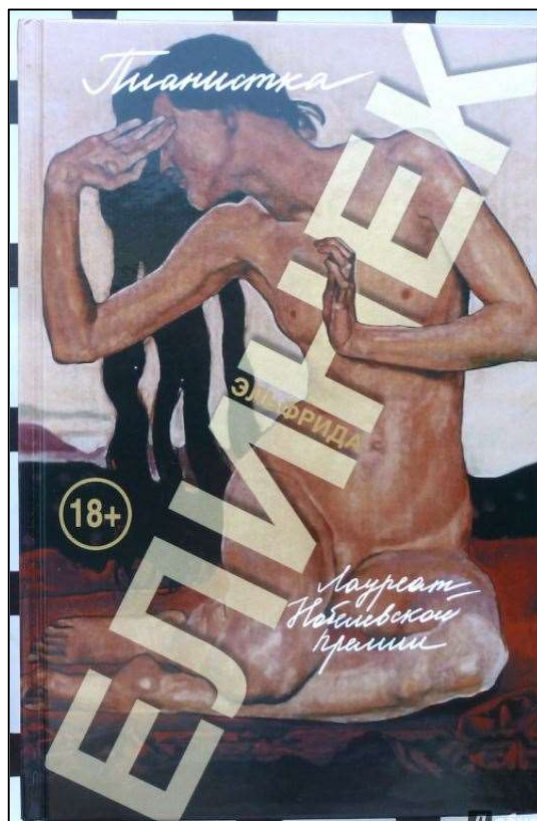


Abbildung 4. Buchcover zum Datensatz PPN 161883679X

⁷⁸ Quelle aller drei Bilder: <https://www.labyrinth.ru/books/481871/> (letzter Zugriff: 01.02.2020).

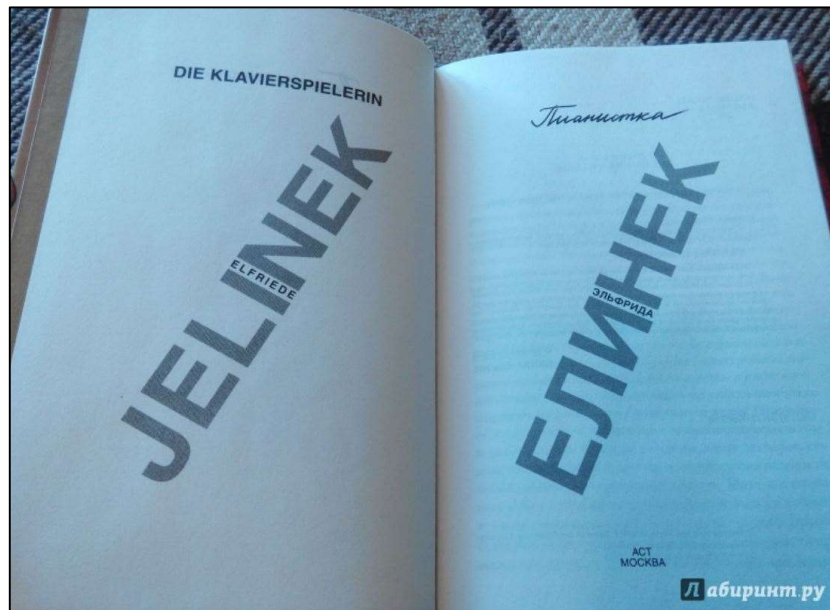


Abbildung 5. Titelblatt zum Datensatz PPN 161883679X

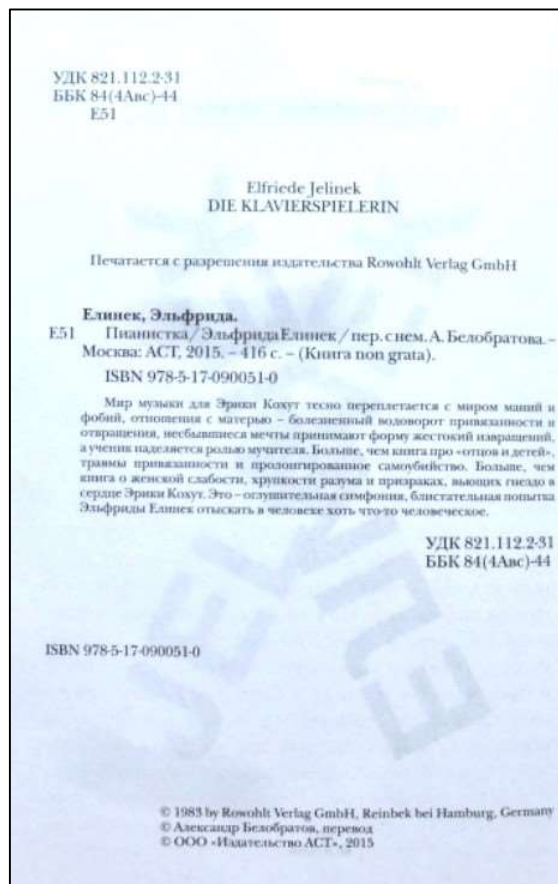


Abbildung 6. Rückseite des Titelblattes zum Datensatz PPN 161883679X

Das Buchcover (Abbildung 4) stellt zwar gemäß RDA 2.2.2.2 keine bevorzugte Informationsquelle dar, wird hier aber trotzdem aufgeführt, um den möglichen Ursprung des Titeltzusatzes zu ermitteln, denn § 115 RAK sah eine größere Vielfalt von möglichen Informationsquellen vor. Das im Online-Buchhandel verfügbare Bild des Titelblattes (Abbildung 5) ist leider am unteren Rand etwas abgeschnitten, vermittelt aber trotzdem einen guten Eindruck und lässt mit hoher Sicherheit vermuten, dass auch auf dem nicht sichtbaren unteren Rand kein Titeltzusatz zu sehen ist. Die Rückseite des Titelblattes (Abbildung 6) enthält detailliertere Informationen zum Buch, vor allem die sogenannte „Annotationskarte“. Diese Annotationskarte soll zur bibliografischen Erfassung der Monografie dienen. Sie ist zwar kein Pflichtelement gemäß dem russischen Verlagsstandard GOST-P 7.0.4-2006 „Angaben zu Druckwerken. Allgemeine Anforderungen und Gestaltungsregeln“⁷⁹, wird aber in der Regel von allen mehr oder weniger seriösen Verlagen auf der Rückseite des Titelblattes gedruckt. Die Annotationskarte muss in Übereinstimmung mit dem Standard GOST 7.51-98 „Katalog- und Karteikarten. Katalogisierung von Publikationen. Inhalt, Datenstruktur und Gestaltung“ gestaltet werden und entspricht zumindest auf den ersten Blick der ISBD-Beschreibung.⁸⁰ Der genaue Vergleich der beiden Beschreibungen würde über den Rahmen dieser Bachelor-Arbeit hinausgehen, aus diesem Grund wird darauf verzichtet. Es lässt sich auf jeden Fall festhalten, dass eine Annotationskarte auf der Rückseite des Titelblattes auch Angaben zum Titeltzusatz enthalten sollte, der jedoch auf der Abbildung 5 nicht zu sehen sind. Somit handelt es sich bei der Erfassung des Titeltzusatzes im Datensatz PPN 161883679X generell um einen Fehler, da der Titeltzusatz im Original gar nicht vorhanden ist.

Die zweite Doppelung wurde im Datensatz PPN 1029266220 festgestellt, welches den folgenden Titel aufweist: „Stepi Orenburgskie\$dk 25-letiju sozdaninija zapovednika "Orenburgskij" : to the 25-anniversary of the "Orenburgskij" reservation establishment“ („Степи Оренбургские\$dk 25-летию создания заповедника "Оренбургский" [...]“, übersetzt „Steppen von Orenburg\$dzum 25-jährigen Jubiläum des Naturschutzparks "Orenburgskij" [...])“. In diesem Fall wurde die Buchstabenkombination „ni“ fälschlicherweise doppelt geschrieben. Außerdem wurde der Buchstabe „v“ fälschlicherweise durch den Buchstaben „b“ ersetzt. Die richtige Schreibweise wäre „Stepi Orenburgskie\$dk 25-letiju sozdaninija zapovednika "Orenburgskij" [...])“.

Wegen der geringen Zahl der festgestellten Doppelungen können keine Muster ihrer Entstehung identifiziert werden. Es wird daher vermutet, dass es sich um reine Flüchtigkeitsfehler handelt, die jedoch einen negativen Einfluss auf das Retrieval haben.

⁷⁹ Vgl. Federal'noe agentstvo po techničeskomu regulirovaniju i metrologii (2006): GOST P 7.0.4.-2006 Izdanija. Vychodnye svedenija, S. 5, Punkt 4.16.

⁸⁰ Vgl. Mežgosudarstvennyj sovet po standartizacii, metrologii i sertifikacii (1998): GOST 7.51-98. Kartočki dlja katalogov i kartotek. Katalogizacija v izdanii. Sostav, struktura dannych i isdatel'skoe oformlenie.

3.2.4.7 Falsche Groß- und Kleinschreibung

Wie der Name der Fehlerart schon andeutet, werden im Rahmen dieser Kategorie Datensätze analysiert, die eine falsche Groß- und Kleinschreibung enthalten.

Die Einstufung, ob die Groß- und Kleinschreibung möglicherweise falsch sei, ist aus mehreren Gründen nicht so einfach im Vergleich zu den vorherigen Fehlerarten. Gemäß Anhang A RDA, Punkt A 1 sollen in Bezug auf die Großschreibung die Regeln der betreffenden Sprache angewendet werden. Gemäß Punkt A 2.1. ist das erste Wort eines Namens dabei groß zu schreiben, sonstige Wörter werden gemäß den Regeln der jeweiligen Sprache groß- bzw. kleingeschrieben. Die entsprechenden Regeln von RAK werden weiter unten behandelt.

Im Russischen werden Wörter außer am Satzanfang generell kleingeschrieben. Die Regeln der russischen Sprache in Bezug auf die Groß- und Kleinschreibung der Namen, die den größten Teil der großzuschreibenden Wörter ausmachen, sind kompliziert. Unterschiedliche Namenskategorien werden einzeln geregelt, dazu existiert noch eine Reihe von Ausnahmen. Verkompliziert wird diese Situation dadurch, dass der Haupttitel auf dem Titelblatt meistens durchgehend großgeschrieben wird. Weder der entsprechende russische Standard im Verlagswesen GOST-P 7.0.4-2006 „Angaben zu Druckwerken. Allgemeine Anforderungen und Gestaltungsregeln“⁸¹ noch entsprechende Handbücher⁸² enthalten eine explizite Regelung dazu, führen jedoch die durchgehende Großschreibung des Haupttitels in sämtlichen Beispielen auf. Eins davon ist in der Abbildung 7 unten aufgeführt.

⁸¹ Vgl. Federal'noe agentstvo po techničeskomu regulirovaniju i metrologii (2006): GOST R 7.0.4.-2006 Izdanija. Vychodnye svedenija. Obščie trebovanija i pravila oformlenija.

⁸² Vgl. Kalinin, S. Ju. (2003): Kak pravil'no oformit' vychodnye svedenija izdanija. Posobie dlja izdatelja; S. 25-31.

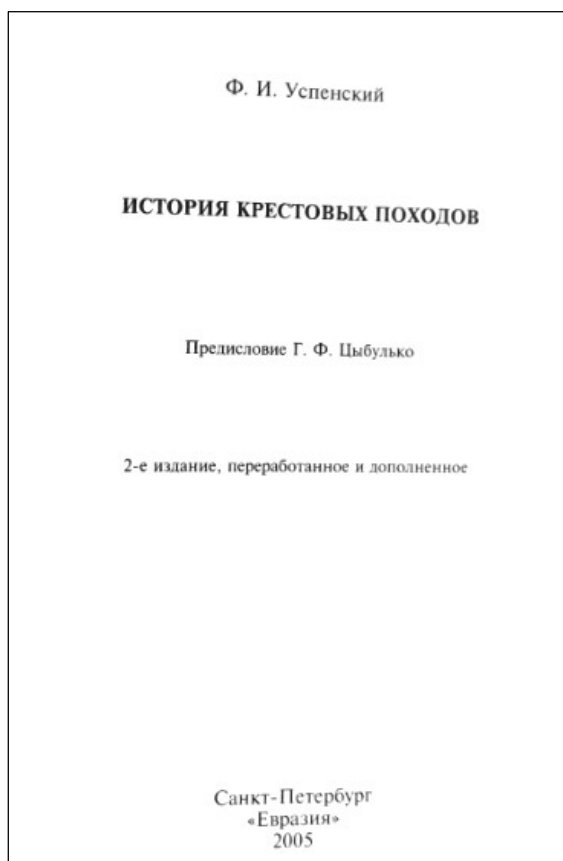


Abbildung 7. Beispiel einer Titelseite nach GOST R 7.0.4.-2006⁸³

Die Abbildung 7 zeigt ein Titelblatt einer Monografie, welches als Beispiel im Standard GOST R 7.0.4.-2006 aufgeführt wird. In der zweiten Zeile ist nach dem Verfassernamen der Haupttitel des Werks „Istorija Krestovych pochodov“ (übersetzt „Geschichte der Kreuzzüge“) durchgehend großgeschrieben und fett markiert.

Dieses Titelblatt ist gleichzeitig ein sehr gutes Beispiel von den Schwierigkeiten, die bei der Katalogisierung russischsprachiger Monografien in Bezug auf die Groß- und Kleinschreibung entstehen. Gemäß § 102 der Regeln der russischen Rechtschreibung und Zeichensetzung aus dem Jahr 1956 war das erste Wort im Namen von historischen Ereignissen, Epochen, historischen Dokumenten, Kunstwerken u. Ä. großzuschreiben. Nachfolgende Wörter wurden kleingeschrieben, außer wenn es sich um Eigennamen handelte. Unter die Großschreibregelung fallen laut § 102 Namen, bestehend aus einem Substantiv (wie Ренессанс (Renessans – Renaissance)), aus einem Substantiv zusammen mit einem Adjektiv, welches von einem Eigennamen gebildet wurde (wie Нантский эдикт (Nantskij édiikt – Edikt von Nantes)) oder aus einer anderen Kombination mit einem

⁸³ Vgl. Federal'noe agentstvo po techničeskomu regulirovaniju i metrologii (2006): GOST R 7.0.4.-2006 Izdanija. Vychodnye svedenija, S. 11.

Anfangsadjektiv⁸⁴ oder einem Zahlwort (Семилетняя война (Semiletnjaja vojna – Siebenjähriger Krieg)). Namen von historischen Ereignissen, Epochen usw., die keine Eigennamen sind, waren dabei kleinzuschreiben (z. B. крестовые походы (krestovye pochody – Kreuzzüge)).⁸⁵ Die Regeln gaben dabei keine Kriterien an, wann es sich bei Namen von historischen Ereignissen, Epochen usw. um Eigennamen handelt und wann nicht, was zusammen mit der aufgeführten Auflistung der unter die Regelung fallenden Fälle die Situation ziemlich verkompliziert. In den aktuellen Regeln der russischen Rechtschreibung wurde diese Regelung gekürzt und vereinfacht und es gilt, dass das erste Wort in Namen von historischen Epochen und Ereignissen, Zeiträumen und Feiertagen großgeschrieben wird (z. B. Крестовые походы (Krestovye pochody)).⁸⁶ Hier gibt es keine detaillierten Regelungen mehr, welche Wörterkombinationen unter die Großschreibregelung fallen sowie keine Unterscheidung von Namen, die keine Eigennamen sind.

Die unverständliche und komplizierte alte Regelung, die noch nicht so lange zurückliegt, sowie ihre Änderung in den aktuellen Rechtschreibungsregeln führten dazu, dass die Großschreibung des Begriffes „Крестовые походы“ (Krestovye pochody) in der Praxis sehr uneinheitlich ist. So lassen sich im OPAC der Russischen Staatsbibliothek in den Aufnahmen aktueller Werke beide Varianten finden – sowohl die Großschreibung als auch die Kleinschreibung.⁸⁷

Dies ist allerdings nur ein Beispiel von den Feinheiten der russischen Großschreibung von Namen. Darüber hinaus gibt es einzelne ausführliche Regeln für die Großschreibung von Eigennamen von Personen, Tieren und mythologischen Wesen, geografischen und administrativen Verwaltungseinheiten, astronomischen Namen, mit Religion verbundenen Namen, Namen von Behörden, Einrichtungen, Organisationen, Vereinen und Parteien und viele andere. In der Praxis werden diese Regeln jedoch oft nicht eingehalten. So kann es zum Beispiel vorkommen, dass eine Konferenz oder eine Einrichtung offiziell jedes Wort des eigenen Namens entgegen der eigentlichen Rechtschreibungsregeln großschreibt. Kein Wunder, dass sogar große russische Bibliotheken gegen diese Verwirrung nicht durchkommen.

In den RAK-Zeiten war die im § 117.6 enthaltene Regelung in Bezug auf die Groß- und Kleinschreibung im Vergleich zu RDA anders:

„Für die Groß- und Kleinschreibung gelten, unabhängig von der Typographie der Vorlage, im allgemeinen die Rechtschreiberegeln der betreffenden Sprache. In Zweifelsfällen richtet man sich nach der Vorlage.“

[...]

⁸⁴ Es handelt sich um eine wörtliche Übersetzung. Im Original wurde das Wort „начальное прилагательное“ verwendet, dessen genaue Begriffserklärung trotz sorgfältiger Recherche nicht ermittelt werden konnte.

⁸⁵ Vgl. § 102, Gramota.ru (o. J.): Pravila russkoj orfografii i puntuacii 1956. Orfografija. Propisnye bukvy.

⁸⁶ Vgl. § 179 in Lopatin, V. V., Hrsg. (2009): Pravila russkoj orfografii i puntuacii.

⁸⁷ Recherche durchgeführt am 13.01.2020.

Mit großen Anfangsbuchstaben werden im allgemeinen geschrieben: das erste Wort eines Sachtitels (auch eines zitierten Sachtitels), einer Angabe einer Unterreihe, Abteilung oder fortlaufenden Beilage; alle Wörter, die nach Punkt, Ausrufe- oder Fragezeichen stehen; außerdem in Eigennamen, Körperschaftsnamen und geographischen Namen alle Wörter außer Artikeln, Präpositionen und Konjunktionen im Innern der Namen.⁸⁸

Somit waren in Eigennamen, Körperschaftsnamen und geographischen Namen praktisch alle Wörter großzuschreiben, und zwar unabhängig von den Regeln der jeweiligen Sprache. Die Auseinandersetzung mit den komplexen Regeln der russischen Groß- und Kleinschreibung war somit nicht erforderlich.

Da für die Stichprobe die Jahrgänge 2011-2015 ausgewählt wurden, stammen viele Datensätze aus den RAK-Zeiten und enthalten somit die Groß- und Kleinschreibung gemäß RAK-Regeln. Aus diesem Grund muss bei der Analyse der möglichen Groß- und Kleinschreibfehler immer berücksichtigt werden, ob der jeweilige Datensatz nach RAK oder nach RDA erfasst worden ist.

So wurden zum Beispiel im Datensatz PPN 1613652496 mit dem Titel „**M**oskovskogo **U**niversiteta XVIII - pervoj četverti XIX veka“ („Студенчество **М**осковского **у**ниверситета XVIII - первой четверти XIX века“, übersetzt „Studierende der Moskauer Universität im 18 – erstem Viertel des 19 Jahrhunderts“) beide Wörter im Namen der Körperschaft großgeschrieben so wie es in RAK vorgesehen war. Die Groß- und Kleinschreibung nach RDA würde sich nach den Regeln der russischen Sprache richten und wäre in diesem Fall „**M**oskovskogo **u**niversiteta XVIII - pervoj četverti XIX veka“. Da die Groß- und Kleinschreibung im Haupttitel den damaligen Regeln entsprach, wurden dieser und ähnliche Datensätze als korrekt eingestuft.

Zweifelsfälle, in denen es (sogar rein theoretisch) Argumente für die im jeweiligen Datensatz gewählte Groß- und Kleinschreibung geben könnte, wurden ebenfalls nicht als Fehler berücksichtigt. Ein Beispiel dafür ist der aus den RAK-Zeiten stammende Datensatz PPN 783671687 mit dem Haupttitel „Archeologičeskie i istoriko-architekturnye issledovanija Mavzoleja Chodži Achmeda Jasavi“ („Археологические и историко-архитектурные исследования мавзолея Ходжи Ахмеда Ясави“, übersetzt „Archäologische und historisch-architektonische Erforschung des Mausoleums von Hodscha Ahmad Yasawi“), in dem das Wort Mausoleum großgeschrieben wurde. Wenn man das Wort als eine beschreibende Angabe zur Art des Baus betrachtet, wäre hier die Kleinschreibung angebracht, da Substantive im Russischen in der Regel kleingeschrieben werden. Andererseits könnte argumentiert werden, dass es sich dabei um einen Individualnamen des Bauwerkes oder sogar um das darin befindende Museum als Körperschaft handelt, wodurch nach den RAK-Regeln eine Großschreibung erforderlich wäre. Das Titelblatt kann in diesem Fall auch keine Klarheit bringen, da der Haupttitel der in

⁸⁸ Deutsche Nationalbibliothek, Hrsg. (2007): Regeln für die alphabetische Katalogisierung in wissenschaftlichen Bibliotheken, S. 21

Kasachstan erschienenen, aber auf Russisch verfassten Monografie gemäß der russischen Verlagstradition durchgehend großgeschrieben wurde (siehe Abbildung 8 unten). Somit wurden dieser und ähnliche Datensätze in Bezug auf die Groß- und Kleinschreibung nicht als fehlerhaft eingestuft.

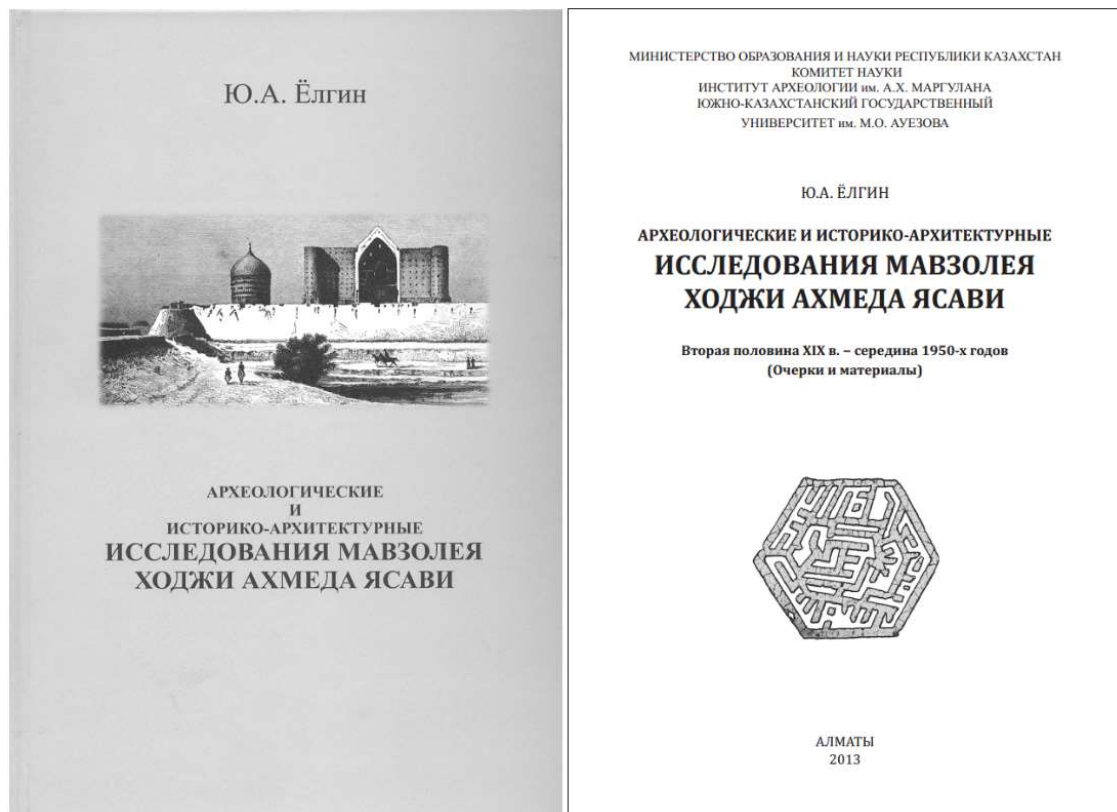


Abbildung 8. Buchcover und Titelblatt zum Datensatz PPN 783671687⁸⁹

Im Rahmen der Datenanalyse wurden ebenfalls Groß- und Kleinschreibfehler festgestellt, die eindeutig sind. In erster Linie handelt es sich um eine inkorrekte Kleinschreibung von Namen sowie um eine inkorrekte Großschreibung von Substantiven. Ein Beispiel für den ersten Fall ist der Datensatz PPN 884609421 mit dem Titel „Novgorodskaja oblast' \$d atlas : karta rossijskoj federacii, obzornye karty Novgorodskoj oblasti, topograficeskaja karta masštaba 1:100 000 Novgorodskoj oblasti, ukazatel' nazvanij naselënych punktov, novgorodskij kreml' i jaroslavovo dvorišče“ („Новгородская область \$d атлас : карта Российской Федерации, обзорные карты Новгородской области, топографическая карта масштаба 1:100 000 Новгородской области, указатель названий населённых пунктов, Новгородский кремль и Ярославово дворище“, übersetzt „Gebiet Nowgorod \$d Atlas : Karte der Russischen Föderation,

⁸⁹ Quelle der Abbildung: <http://tehne.com/library/yolgin-yu-arheologicheskie-i-istoriko-arhitekturnye-issledovaniya-mavzoleya-hodzhi-ahmeda-yasavi-vtoraya-polovina-xix-v-seredina-1950-h-godov-almaty-2013> (letzter Zugriff: 26.02.2020)

Übersichtskarten des Gebiets Nowgorod, topografische Karte des Gebiets Nowgorod im Maßstab 1:100 000, Verzeichnis von Ortschaften, Nowgoroder Kreml und Jaroslaw-Hof“). Der Datensatz wurde nach RDA katalogisiert, die Namen Russische Föderation, Nowgoroder Kreml und Jaroslaw-Hof wurden dabei kleingeschrieben. Die richtige Schreibweise wäre „Novgorodskaja oblast'**\$d**atlas : karta **R**ossijskoj **F**ederacii, obzornye karty Novgorodskoj oblasti, topografičeskaja karta masštaba 1:100 000 Novgorodskoj oblasti, ukazatel' nazvanij naselënných punktov, **N**ovgorodskij kremli i **J**aroslavovo dvoriščë“. Somit liegen in diesem Fall drei Groß- und Kleinschreibfehler vor.

Ein Beispiel für die inkorrekte Großschreibung von Substantiven ist der RAK-Datensatz PPN 1615742255 mit dem Titel „Kniga-trener ubeždenija, umenija vlijat' na **L**judej i preobretat' družej; intensivnyj kurs“ („Книга-тренер убеждения, умения влиять на людей и приобретать друзей; интенсивный курс“, übersetzt „Übungsbuch, wie man andere Menschen überzeugt und beeinflusst sowie Freunde gewinnt; Intensivkurs“). Hier wurde das russische Substantiv „Menschen“ großgeschrieben, wozu es keinen Anlass gibt. Außerdem handelt es sich beim Wort „Intensivkurs“ eher um einen Titelzusatz, da es lediglich weiterführende Informationen zur Art des Werkes liefert. Die korrekte Schreibweise wäre in diesem Fall „Kniga-trener ubeždenija, umenija vlijat' na **I**judej i preobretat' družej**\$d**intensivnyj kurs“. Da die Erfassung eines Titelzusatzes nicht mit Transliteration verbunden ist, wurde in diesem Beispiel nur ein Groß- und Kleinschreibfehler berücksichtigt.

Eine inkorrekte Großschreibung von Substantiven kommt häufig im ersten Wort des Titelzusatzes vor. Ein Beispiel dafür ist der Datensatz PPN 1608252787 mit dem folgenden Feld 4000 „Aptekarša **\$dR**oman\$ Ingrid Noll. Übers. von G. Ševčenko“ („Аптекарьша**\$d**роман [...]", übersetzt „Die Apothekerin **\$d**Roman\$ Ingrid Noll. Übers. von G. Ševčenko“). Das Substantiv „Roman“ wird in der russischen Sprache kleingeschrieben, dabei sehen weder RDA noch RAK eine Großschreibung des Titelzusatzes vor (vgl. 1.7 RDA und § 117.6 RAK). Hier lässt es sich schwer beurteilen, ob die falsche Großschreibung mit inkorrektur Transliteration wegen mangelnder Sprachkenntnisse oder mit einem Fehler wegen mangelnder Regelwerkskenntnisse verbunden ist. Im Rahmen dieser Bachelor-Arbeit wurde entschieden, derartige Fehler als Groß- und Kleinschreibfehler zu berücksichtigen, denn wie die Abbildung 8 zeigt, gibt es durchaus Fälle, wenn das erste Wort des Titelzusatzes auf dem Titelblatt großgeschrieben wird und das bei mangelnden Sprachkenntnissen auch beim Katalogisieren fälschlicherweise übernommen werden könnte. Außerdem wurden im Feld 4000 des Beispiels fälschlicherweise ein Leerzeichen vor dem Code „\$d“ hinzugefügt sowie das „h“ im Code „\$h“ für die Verantwortlichkeitsangabe ausgelassen. Die Verantwortlichkeitsangabe wird im Rahmen dieser Bachelor-Arbeit nicht behandelt, aber auch hier lässt sich kurz erwähnen, dass das Vorkommen des deutschen Wortes „Übersetzt“ auf dem Titelblatt einer russischsprachigen Monografie äußerst unwahrscheinlich ist. Womöglich handelt es sich dabei um eine Übersetzung der Verantwortlichkeitsangabe anstatt ihrer Transliteration. Da Leerzeichen und Codes nicht mit Transliteration verbunden sind, wurde in diesem Beispiel ebenfalls nur ein Groß- und Kleinschreibfehler im Titelzusatz berücksichtigt. Die korrekte Schreibweise wäre in diesem Fall „Aptekarša**\$d**roman“.

Auf Grundlage der beschriebenen Überlegungen zur Berücksichtigung der Groß- und Kleinschreibfehler wurden im Rahmen der Datenanalyse 23 Datensätze festgestellt, die insgesamt 28 Groß- und Kleinschreibfehler enthielten. Somit beträgt der Anteil der Datensätze, die Groß- und Kleinschreibfehler enthalten, an der Gesamtzahl der geprüften Datensätze 28,8 %. Der Anteil der Groß- und Kleinschreibfehler an der Gesamtfehlerzahl liegt bei 25,5 %.

Auf die Aufführung der einzelnen Beispiele wird in Bezug auf die Groß- und Kleinschreibfehler verzichtet, da in den vorangehenden Überlegungen einige Beispiele bereits vorgebracht worden sind.

Die einzelnen in der Stichprobe festgestellten Groß- und Kleinschreibfehler wurden nach Fehlerarten analysiert und in Tabelle 16 zusammengefasst.

Tabelle 16. Aufschlüsselung der Groß- und Kleinschreibfehler

Fehlerart	Anzahl der Fehler in der Stichprobe
Kleinschreibung von Namen	8
Großschreibung von Substantiven	20
darunter im ersten Wort des Titelsatzes	11
darunter im Inneren des Haupttitels oder des Titelsatzes	9
Insgesamt	28

Die Aufschlüsselung der Groß- und Kleinschreibfehler zeigt, dass es sich bei mehr als zwei Drittel der Fehler (20 Fehler) um die falsche Großschreibung der Substantive handelt. Falsche Kleinschreibung der Namen kommt dabei dementsprechend in weniger als einem Drittel der Fälle vor (8 Fehler). Bei der falschen Großschreibung der Substantive handelt es sich in ca. einer Hälfte der Fälle (9 Fehler) um eine falsche Großschreibung im Inneren des Haupttitels oder des Titelsatzes und in der anderen Hälfte (11 Fehler) um die falsche Großschreibung im ersten Wort des Titelsatzes.

Da die Groß- und Kleinschreibung bei der Indexierung nicht berücksichtigt wird, hat sie dementsprechend keinen Einfluss auf das Retrieval.

3.2.4.8 Falscher Transliterationsstandard

Den in vorherigen Punkten beschriebenen Fehlerarten lag eine falsche Anwendung des für die Formalerschließung geltenden Transliterationsstandards DIN 1460 zugrunde. Während der Datenanalyse wurden jedoch auch Datensätze festgestellt, in denen für die Transliteration nicht DIN 1460, sondern ein ganz anderer Transliterationsstandard

verwendet wurde. Insgesamt konnten 8 solcher Datensätze festgestellt werden. Da diese Fehlerart sich auf die Schreibweise des jeweiligen transliterierten Felds als Ganzes bezieht und falsche Buchstaben und Buchstabenkombinationen daher auch mehrmals vorkommen, wurde in Bezug auf diese Fehlerart jeder einzelne Datensatz und nicht jeder falsche Buchstabe oder Buchstabenkombination einzeln als ein Fehler berücksichtigt. Die Anzahl der festgestellten Fehler ist somit gleich der Zahl der diese Fehlerart enthaltenden Datensätze. Somit liegt der Anteil der nach einem falschen Standard transliterierten Datensätze an der Gesamtzahl der fehlerhaften Datensätze bei 10 % und der Anteil dieser Fehlerart an der Gesamtfehlerzahl bei 7,3 %.

Bei den fehlerhaften Datensätzen lässt sich ein unterschiedliches Ausmaß feststellen, zu dem ein falscher Transliterationsstandard angewendet wurde. In manchen Datensätzen wurde ein falscher Transliterationsstandard im Haupttitel und Titeltzusatz durchgehend verwendet. Ein Beispiel dafür ist der Datensatz PPN 161035138X mit dem folgenden Titel: „Ekaterina II\$**d**put' k tronu : k 250-letiiu voss**sh**estviia na prestol“ („Екатерина II\$**d**путь к трону : к 250-летию вос**sh**ествия на престол“, übersetzt: „Ekaterina II\$**d**der Weg zum Thron : anlässlich des 250-jährigen Jubiläums der Thronbesteigung“). Die abweichende Transliteration der russischen Buchstaben „ю“, „ш“ und „я“ lässt erkennen, dass im Datensatz der Transliterationsstandard der Library of Congress und der American Library Association (siehe Tabelle 1 im Punkt 2.1.3.2) angewendet wurde. Die richtige Transliteration nach DIN 1460 wäre „Ekaterina II\$**d**put' k tronu : k 250-letiju voss**š**estvjia na prestol“.

Es gab ebenfalls Datensätze, in denen zwei Transliterationsstandards gleichzeitig angewendet wurden, wie zum Beispiel der Datensatz PPN 78425219X mit dem Titel „Kak duch naš gorestnyj živu**ch** ..."\$**d**Stat'i, ésse, vospominanija, pis'ma, zametki iz literaturnogo dnevnika, stichi“ („Как дух наш горестный живу**ч**..."\$**d**статьи, эссе, воспоминания, письма, заметки из литературного дневника, стихи“, übersetzt „Wie zäh ist unser wehmütiger Geist..."\$**d**Artikel, Essays, Erinnerungen, Briefe, Tagebuchnotizen, Gedichte“). Die Transliteration des russischen Buchstabens „ч“ am Ende des Haupttitels entspricht dabei dem Transliterationsstandard der Library of Congress und der American Library Association sowie dem Teil B des russischen Standards GOST 7.79-2000 (ISO 9-95) (siehe Tabelle 1 im Punkt 2.1.3.2). Alle restlichen Buchstaben wurden nach DIN 1460 transliteriert. Die richtige Transliteration des Titels nach DIN 1460 wäre „Kak duch naš gorestnyj živu**č** ..."\$**d**stat'i, ésse, vospominanija, pis'ma, zametki iz literaturnogo dnevnika, stichi“. Im Feld 4212 des Datensatzes wurde zusätzlich als abweichender Titel die alternative Variante "Kak dukh nash gorestnyï zhivuch--" erfasst, bei der der Transliterationsstandard der Library of Congress und der American Library Association verwendet wurde. Dies könnte der Grund sein, warum eine Buchstabenkombination aus dem abweichenden Standard auch im Haupttitel versehentlich verwendet worden ist.

Auch in den restlichen Datensätzen wurden fälschlicherweise Buchstabenkombinationen aus dem amerikanischen Transliterationsstandard verwendet. Dies könnte darauf

hindeuten, dass während der Katalogisierung amerikanische Fremddaten als Grundlage hinzugezogen wurden. Dies wurde jedoch im Rahmen der Analyse nicht überprüft.

Die Anwendung eines falschen Transliterationsstandards kann dabei unterschiedliche Auswirkungen auf den jeweiligen Datensatz haben und hängt davon ab, welche Wörter im Titelfeld vorkommen. Im Datensatz PPN 787328073 mit dem Titel „Infrastrukturizma\$du**ch**ebnoe posobie“ („Инфраструктура туризма\$**д**учебное пособие“, übersetzt „Turismusinfrastruktur**д**Lehrbuch“) wurde der russische Buchstabe „ч“ im Titelnachsatz nach dem Transliterationsstandard der Library of Congress und der American Library Association bzw. Teil B des russischen Standards GOST 7.79-2000 (ISO 9-95) transliteriert. Die richtige Transliteration nach DIN 1460 wäre „Infrastrukturizma\$du**č**ebnoe posobie“. Bei allen anderen Buchstaben des Haupttitels und des Titelnachsatzes spielt die Wahl des Transliterationsstandards allerdings keine Rolle, denn das Ergebnis bliebe bei allen in Tabelle 1 im Punkt 2.1.3.2 aufgeführten Standards gleich. Somit wäre der kleine Transliterationsfehler in diesem Datensatz vermutlich nicht retrievalrelevant, denn es ist eher unwahrscheinlich, dass nach dem ohnehin nicht viel-sagenden Titelnachsatz recherchiert wird. Generell sind die Fehler dieser Art jedoch retrievalrelevant.

Im ersten in diesem Unterpunkt aufgeführten Beispiel (PPN 161035138X) hat die Wahl des Transliterationsstandards ebenfalls keine Auswirkungen auf die Schreibweise des Haupttitels. Hier könnte es allerdings Schwierigkeiten geben, wenn genau diese Monografie gesucht wird. Da Ekaterina die Große eine bedeutende Persönlichkeit war und über sie viel geschrieben wird, wäre im Falle eines Known-Item-Searches unter anderem die Auswahl des Haupttitels sowie des Titelnachsatzes oder seiner Teile für die Recherche denkbar. In so einer Situation wäre der Unterschied in der Schreibweise und somit in der Indexierung womöglich zu groß und somit auch retrievalrelevant. Diese Hypothese kann jedoch nur durch eine praktische Untersuchung bestätigt werden, die im Rahmen dieser Bachelor-Arbeit nicht vorgesehen ist.

3.2.4.9 Falsche diakritische Zeichen

Wie bereits angesprochen, war die Fehlerart „falsche diakritische Zeichen“ bei der Formulierung der möglichen Fehlerarten im Punkt 3.1.2 ursprünglich nicht vorgesehen. Während der Datenanalyse wurden jedoch Datensätze festgestellt, bei denen der Grundbuchstabe richtig war, jedoch falsche diakritische Zeichen verwendet worden sind. Diese Fehler wurden anfangs als Ersetzungen erfasst. Nach dem näheren Betrachten wurde jedoch entschieden, sie in eine eigene Fehlerart auszugliedern, da es sich wegen dem gleichbleibenden Grundbuchstaben nicht um eine reine Ersetzung handelt.

Insgesamt wurden drei Datensätze festgestellt, die jeweils einen Fehler dieser Art enthielten. Somit liegt der Anteil der Datensätze mit falschen diakritischen Zeichen bei 3,8 % der Gesamtzahl der fehlerhaften Datensätze in der Stichprobe. Der Anteil dieser Fehlerart an der Gesamtzahl der festgestellten Fehler beträgt 2,7 %.

Bei allen Fehlern handelt es sich um die Verwendung eines Breves (rundliches diakritisches Zeichen über einem Buchstaben) anstatt eines Hatscheks (eckiges diakritisches Zeichen), welches durch die DIN 1460 vorgesehen ist. Betroffen waren die Buchstaben „š“ (russisches „ш“), „ž“ (russisches „з“) und „šč“ (russisches „щ“). Die entsprechenden Datensätze werden in Tabelle 17 unten aufgeführt.

Tabelle 17. Verwendung falscher diakritischer Zeichen

PPN	Fehlerhafte Schreibweise	Korrekte Schreibweise	Schreibweise in Originalschrift
717065324	Ugolovnoe pravo v rešenijach Konstitucionnogo Suda Rossijsko <i>i</i> Federacii \$dmonografija	Ugolovnoe pravo v rešenijach Konstitucionnogo Suda Rossijsko <i>j</i> Federacii \$dmonografija	Уголовное право в решен <i>и</i> ях Конституционного суда Российско <i>й</i> Федерации \$dмонография
160458954X	Ljubov' i logika prostranctva \$dskul'ptura, živopis'	Ljubov' i logika prostranstva \$dskul'ptura, živopis'	Любовь и логика простран <i>с</i> тва \$dскульптура и живопис <i>ь</i>
884609421	Novgorodskaja oblast'\$datlas : karta rossijskoj federacii, obzornye karty Novgorodskoj oblasti, topografičeskaja karta masštaba 1:100 000 Novgorodskoj oblasti, ukazatel' nazvanij naselënných punktov, novgorodskij kreml' i jaroslavovo dvorišče	Novgorodskaja oblast'\$datlas : karta Rossijskoj Federacii, obzornye karty Novgorodskoj oblasti, topografičeskaja karta masštaba 1:100 000 Novgorodskoj oblasti, ukazatel' nazvanij naselënných punktov, Novgorodskij kreml' i Jaroslavovo dvorišče	Новгородская область\$датлас : карта Российской Федерации, обзорные карты Новгородской области, топографическая карта масштаба 1:100 000 Новгородской области, указатель названий населённых пунктов, Новгородский кремль и Ярославово дворище

In allen drei Fällen der Verwendung falscher diakritischer Zeichen war ein jeweils anderer Grundbuchstabe betroffen. Aus diesem Grund und wegen der insgesamt niedrigen Fehlerzahl kann für diese Fehlerart kein Muster identifiziert werden. Es kann auch schwer erklärt werden, warum es in der Praxis zu einer solchen Fehlerart kommt, außer dass die beiden diakritischen Zeichen ähnlich aussehen und die Verwechslungsgefahr groß ist. Es muss zudem erwähnt werden, dass derartige Fehler wegen der minimalen graphischen Unterschiede zwischen einem Breve und einem Hatschek sich sehr schwer feststellen lassen. Es könnte sein, dass die Anzahl derartiger Fehler höher ist.

Da diakritische Zeichen bei der Indexierung jedoch wegfallen und nur der Grundbuchstabe indexiert wird, ist diese Fehlerart nicht retrievalrelevant.

3.2.4.10 Sonstige Auffälligkeiten

Während der Analyse wurden zusätzlich zu Transliterationsfehlern weitere Auffälligkeiten festgestellt. Da sie nicht transliterationsbedingt sind, wurden sie nicht als Fehler im Rahmen dieser Bachelor-Arbeit eingestuft und werden daher als Auffälligkeiten bezeichnet. Sie sollen aber trotzdem am Rande erwähnt werden und werden im Folgenden kurz beschrieben.

Als Erstes lassen sich Auffälligkeiten im Titelnzusatz unterscheiden. Beispiele für diese Kategorie sind in Tabelle 18 aufgeführt.

Tabelle 18. Sonstige Auffälligkeiten, Teil 1

PPN	Fehlerhafte Schreibweise	Korrekte Schreibweise	Schreibweise in Originalschrift
825563232	Simmetrija i asimetrija \$dgrammatičeskoj kategorii roda i semantičeskoj kategorii pola v slavjanskih jazykach	Simmetrija i asimetrija grammatičeskoj kategorii roda i semantičeskoj kategorii pola v slavjanskih jazykach	Симметрия и асимметрия грамматической категории рода и семантической категории пола в славянских языках
1615742255	Kniga-trener ubeždenija, umenija vlijat' na Ljudej i preobretat' druzej; intensivnyj kurs	Kniga-trener ubeždenija, umenija vlijat' na ljudej i preobretat' druzej \$dintensivnyj kurs	Книга-тренер убеждения, умения влиять на людей и приобретать друзей \$динтенсивный курс

PPN	Fehlerhafte Schreibweise	Korrekte Schreibweise	Schreibweise in Originalschrift
815431600	Ob istine, žizni i povenii \$d[(orys tilinde)]	Ob istine, žizni i povenii	Об истине, жизни и поведении
1603211004	Nevrozy \$dklinika, profilaktika i lečenje(gaz)	Nevrozy \$dklinika, profilaktika i lečenje	Неврозы \$dклиника, профилактика и лечение

Wie die in Tabelle 18 aufgeführten Beispiele zeigen, handelt es sich dabei um Auffälligkeiten unterschiedlicher Natur. Im ersten Beispiel PPN 825563232 wurde der Haupttitel fälschlicherweise als Haupttitel und Titeltzusatz aufgenommen. Übersetzt sieht der Titel im Datensatz folgendermaßen aus: „Symmetrie und Asymmetrie \$d des grammatischen und des semantischen Geschlechts in slawischen Sprachen“. Dabei handelt es sich jedoch sinngemäß um einen Haupttitel, der im Datensatz durch den Code „\$d“ in der Mitte unterbrochen wird. Es könnte vermutet werden, dass die Gestaltung des Titelblattes verwirrend gewesen sein könnte, indem Teile des Haupttitels in unterschiedlichen Schriftgrößen oder auf getrennten Zeilen geschrieben waren. Ohne Sprachkenntnisse wäre der Haupttitel in so einem Fall schwierig zu erkennen.

Das zweite Beispiel PPN 1615742255 zeigt einen gegenteiligen Fall. Hier wurden Haupttitel und Titeltzusatz „intensivnyj kurs“ (übersetzt „Intensivkurs“) im Haupttitelfeld erfasst, obwohl die Bezeichnung „Intensivkurs“ ins Unterfeld Titeltzusatz gehört. Auch hier könnte der Grund in fehlenden Sprachkenntnissen liegen.

In den letzten zwei Beispielen liegt das Problem ebenfalls beim Titeltzusatz, der in beiden Fällen jedoch keinen Sinn ergibt. Der Titeltzusatz „\$d[(orys tilinde)]“ im Datensatz PPN 815431600 ist weder in den russischen OPACs noch auf dem Buchcover, dem Titelblatt oder auf der Rückseite des Titelblattes⁹⁰ zu finden. Darüber hinaus existieren die beiden Wörter in der russischen Sprache nicht. Eine Internet-Recherche ergab, dass „орыс тілінде“ in der kasachischen Sprache „auf Russisch“ bedeutet. Somit ergibt der Titeltzusatz keinen Sinn, sein Ursprung ist unklar.

Der Datensatz PPN 1603211004 weist den folgenden Titel auf: „Nevrozy \$d klinika, profilaktika i lečenje(gaz)“ (übersetzt „Neurosen \$d Symptome, Vorbeugung und Behandlung (Gas)“). Weder in den russischen OPACs noch auf dem Buchcover⁹¹ ist das Wort „Gas“ am Ende des Titeltzusatzes vorhanden. Beim Wort „gaz“ handelt es sich zwar um ein existierendes russisches Wort, welches im Kontext eines Buches über Neurosen jedoch keinen Sinn ergibt. Der Ursprung des Titeltzusatzes ist ebenfalls unklar. Außerdem hätte

⁹⁰ Entsprechende Abbildungen sind im Online-Buchhandel verfügbar. URL: <https://book24.ua/ua/product/ob-istine-zhizni-i-povedenii-zbm/> (letzter Zugriff: 18.02.2020)

⁹¹ Im Online-Buchhandel ist in Bezug auf die 4. Auflage nur das Buchcover verfügbar. URL: <https://www.bookvoed.ru/book?id=616060> (letzter Zugriff: 18.02.2020)

vor der ersten Klammer ein Leerzeichen stehen sollen. Ein Leerzeichen fehlt ebenfalls zwischen den ersten zwei Wörtern des Titelzusatzes.

Beispiele für die zweite Gruppe der Auffälligkeiten werden unten in Tabelle 19 aufgeführt.

Tabelle 19. Sonstige Auffälligkeiten, Teil 2

PPN	Fehlerhafte Schreibweise	Korrekte Schreibweise	Schreibweise in Originalschrift
1620227169	Graždanskaja vojna na Ukraine \$d (1917 - 1920 gg)	Graždanskaja vojna v Ukraine \$d (1917 - 1920 gg.)	Гражданская война в Украине \$d (1917-1920 гг.)
781223512	Ėtnokalendar' Kalužskoj oblasti \$d kalendar' pamjatnich let Kalužskoj oblasti - 2013	Ėtnokalendar' Kalužskoj oblasti \$d kalendar' pamjatnych dat Kalužskoj oblasti - 2013	Этнокалендарь Калужской области \$d календарь памятных дат Калужской области - 2013
817178643	Problemy archeologii, ėtnografii, antropologii Sibiri i sopredel'nych territorij, tom 19 \$d materialy godovoj Sessii Instituta Archeologii i Ėtnografii SO RAN 2013 goda	Problemy archeologii, ėtnografii, antropologii Sibiri i sopredel'nych territorij \$d materialy itogovoj sessii Instituta Archeologii i Ėtnografii SO RAN 2013 god	Проблемы археологии, этнографии, антропологии Сибири и сопредельных территорий \$d материалы итоговой сессии Института археологии и этнографии СО РАН 2013 год

Diese Gruppe der Auffälligkeiten ist äußerst interessant, den hier wurden Wörter aus dem Titel durch ein falsches, aber sinngemäß passendes russisches Wort ersetzt.

Im ersten Beispiel PPN 1620227169 wurde der Haupttitel als „Graždanskaja vojna **na** Ukraine“ (übersetzt „Bürgerkrieg in der Ukraine“) erfasst. Die russischen OPACs sowie das Buchcover⁹² führen jedoch die Schreibweise „Graždanskaja vojna **v** Ukraine“ auf. Der Unterschied liegt dabei in einer Präposition, ist aber nicht zufällig. In der russischen und ukrainischen Gesellschaft kommt es seit Jahren zu heftigen politisch geprägten und

⁹² Im Online-Buchhandel ist nur das Buchcover verfügbar.
URL: <https://bookmix.ru/book.phtml?id=617430> (letzter Zugriff: 18.02.2020)

emotional geführten Diskussionen, welche Präposition im Russischen in Verbindung mit dem Landesnamen „Ukraine“ verwendet werden soll. Die klassische literarische Tradition war über viele Jahrzehnte „na Ukraine“ (wörtlich übersetzt „auf der Ukraine“). Die Gegner dieser literarischen Tradition argumentieren, dass die Präposition „na“ im Russischen analog zum deutschen „auf“ in Bezug auf Inseln oder höchstens Inselstaaten verwendet werde. Außerdem klinge die Kombination „na Ukraine“ ähnlich wie „na okraïne“, was „am Rande“ bedeutet. Beides wird von einem Teil der russischsprachigen ukrainischen Bevölkerung sowie von einem Teil der aus der Ukraine stammenden russischen Bevölkerung als beleidigend empfunden und es wird auf der Bezeichnung „v Ukraine“ (wörtlich „in der Ukraine“) bestanden.

Im zweiten Beispiel PPN 781223512 wurde der Titel „Ėtnokalendar' Kalužskoj oblasti\$dkalendar' pamjatnich let Kalužskoj oblasti – 2013“ (übersetzt „Ethnokalendar des Gebiets Kaluga\$dkalendar mit Gedenk Jahren des Gebiets Kaluga“) erfasst. Hier wurde unter anderem das Wort „let“ („Gedenkjahre“) anstatt „dat“ („Gedenktage“) aufgenommen, denn die russischen OPACs führen den Titel wie folgt auf: „Ėtnokalendar' Kalužskoj oblasti\$dkalendar' pamjatnych dat Kalužskoj oblasti – 2013“. Diese zwei Wörter haben jedoch semantische Ähnlichkeiten, da sie Begriffe in Bezug auf die Zeiterfassung beschreiben.

Der letzte Datensatz PPN 817178643 weist den Titel „Problemy archeologii, étnografii, antropologii Sibiri i sopredel'nych territorij, tom 19\$dmaterialy godovoj Sessii Instituta Archeologii i Étnografii SO RAN 2013 goda“ (übersetzt „Probleme der Archäologie, Ethnografie und Anthropologie von Sibirien und benachbarten Gebieten, Band 19\$dmaterialien der jährlichen Versammlung des Institutes für Archäologie und Ethnografie der Sibirischen Abteilung der Russischen Akademie der Wissenschaften im Jahr 2013“). Hier liegen mehrere Fehler vor (Bandbezeichnung im Haupttitel, falsche Groß- und Kleinschreibung, Einfügen). Im Sinne dieses Unterpunktes ist der Vertausch des Wortes „godovoj“ („jährlich“) interessant. Im Titelblatt⁹³ steht an dieser Stelle das Wort „itogovoj“ (übersetzt „abschließend“), somit handelt es sich nicht um eine jährliche Versammlung, sondern über eine Abschlussversammlung. Sinngemäß ist das fälschlicherweise erfasste Wort „godovoj“ jedoch durchaus passend.

Alle drei aufgeführten Beispiele haben gemeinsam, dass im Titel ein Wort vertauscht wurde, die Vertauschung jedoch nicht zufällig ist und das fälschlicherweise verwendete Wort durchaus Sinn ergibt. Daraus lässt sich schließen, dass die entsprechenden Datensätze von Personen mit guten Russischkenntnissen angelegt wurden.

Im Gegensatz zur zweiten Gruppe der Auffälligkeiten erweckt die letzte dritte Gruppe den Anschein, zufällig zu sein. Beispiele dafür werden in Tabelle 20 aufgeführt.

⁹³ Quelle: https://archaeology.nsc.ru/wp-content/uploads/2018/05/ses_2013.pdf (letzter Zugriff: 18.02.2020)

Tabelle 20. Sonstige Auffälligkeiten, Teil 3

PPN	Fehlerhafte Schreibweise	Korrekte Schreibweise	Schreibweise in Originalschrift
788447254	Informacionno-telekommunikacionn ye technologii i matematičeskoe modelirovanie vysokotechnologičny ch sistem\$dmaterialy vserossijskoj konferencii s meždunarodnym učastiem, Moskva, RUDN, 22 - 27 aprilja 2014 goda	Informacionno-telekommunikacionn ye technologii i matematičeskoe modelirovanie vysokotechnologičny ch sistem\$dmaterialy vserossijskoj konferencii s meždunarodnym učastiem, Moskva, RUDN, 22- 25 aprilja 2014 goda	Информационно-телекоммуникационн ые технологии и математическое моделирование высокотехнологичных систем\$dmaterialы всероссийской конференции с международным участием, Москва, РУДН, 22– 25 апреля 2014 года
1608617386	Russkaja proza konca XX - načala veka\$dOsnovnye tendencii; učebnoe posobie dlja studentov-filologov	Russkaja proza konca XX - načala XXI veka\$dOsnovnye tendencii; učebnoe posobie dlja studentov-filologov	Русская проза конца XX - начала XXI века\$dосновные тенденции; учебное пособие для студентов-филологов
126840121	@Muzykal'noe ispolnitel'stvo i sovremennost'	M uzykal'noe ispolnitel'stvo i sovremennost'	Музыкальное исполнительство и современность
316425176	@Pravda Stolypina\$d[al'manach 1]	P ravda Stolypina\$d[al'manach 1]	Правда Столыпина\$dальман ах

Die aufgeführten Beispiele zeigen, dass es sich eher um gewöhnliche Flüchtigkeitsfehler handelt. Im ersten Beispiel PPN 788447254 wurde eine Zahl falsch geschrieben, im zweiten Beispiel PPN 1608617386 eine Zahl ausgelassen. In den beiden letzten Beispielen wurde ein Klammeraffe vor dem Haupttitel erfasst, der normalerweise zur Markierung der nicht zu indexierenden Wörter dient. In den aufgeführten Beispielen handelt es sich jedoch um normale Wörter. Somit sind der Zweck und der Ursprung des Klammeraffens unklar. Hier könnte es sich eventuell um einen Tippfehler handeln.

Da die beschriebenen Auffälligkeiten nicht transliterationsbedingt sind und somit nicht zum eigentlichen Thema dieser Bachelor-Arbeit gehören, wird es auch nicht angestrebt, daraus weitgehende Schlüsse zu ziehen. Es kann jedoch festgehalten werden, dass

fehlende Sprachkenntnisse nicht nur einer der Gründe für mögliche Transliterationsfehler sein können, sondern auch zu Schwierigkeiten beim Abgrenzen zwischen dem Haupttitel und dem Titeltzusatz führen.

3.3 Beantwortung der Forschungsfragen und Zusammenfassung der Analyseergebnisse

Im Punkt 3.1.1 wurden Forschungsfragen formuliert. Im Folgenden werden die formulierten Forschungsfragen zum besseren Verständnis wiederholt aufgeführt und anhand der gewonnenen Analyseergebnisse beantwortet.

- **Welche Transliterationsfehler entstehen bei der Formalerschließung russischsprachiger Monografien? Lassen sich diese Fehler nach Fehlerarten kategorisieren?**

Bei der Formalerschließung russischsprachiger Monografien entsteht eine Reihe von unterschiedlichen Transliterationsfehlern. Diese Fehler lassen sich gut nach Fehlerarten unterscheiden. In der untersuchten Stichprobe lassen sich neun Fehlerarten auseinanderhalten, und zwar Vertauschung, Auslassung, Weglassung der diakritischen Zeichen, Einfügen, Ersetzung, Doppelung, falsche Groß- und Kleinschreibung, Verwendung eines falschen Transliterationsstandards sowie falscher diakritischer Zeichen.

- **Wie viele Datensätze der Stichprobe (Prozentzahl) enthalten Fehler in den analysierten Feldern?**

Von insgesamt 552 geprüften Datensätzen enthielten 80 Datensätze Transliterationsfehler. Der Anteil der fehlerhaften Datensätze an der Gesamtzahl der geprüften Datensätze liegt somit bei 14,5 %. Da einige dieser Datensätze mehrere Fehler gleichzeitig enthielten, belief sich die Gesamtzahl der festgestellten Transliterationsfehler auf 110.

- **Wie oft kommen einzelne Fehlerarten vor? Können dabei Muster identifiziert werden?**

Die Anzahl der Fehler ist je nach Fehlerart sehr unterschiedlich.

Mit 30 festgestellten Fehlern ist Ersetzung die häufigste Fehlerart, die in der Stichprobe vorkommt. Ihr Anteil an der Gesamtfehlerzahl liegt bei 27,3 %. Insgesamt kamen in der Stichprobe 21 Kombinationen vor, wenn ein Buchstabe durch einen anderen ersetzt worden ist, 16 davon waren einmalig. Obwohl in Bezug auf die Ersetzung bestimmter Buch-

staben sprachliche Hintergründe vermutet werden können, lassen sich wegen der niedrigen absoluten Zahlen keine eindeutigen Muster feststellen. Ersetzung gehört dabei zu den Fehlern, die retrievalrelevant sind.

Die Fehlerart „falsche Groß- und Kleinschreibung“ wies sich als komplex aus in Bezug auf die Beurteilung, ob ein Datensatz korrekt ist oder einen Fehler enthält. Dies ist zum einen auf die unterschiedlichen Regelungen von RDA und RAK und zum anderen auf die komplexen russischen Rechtschreibregeln zurückzuführen. Aus diesem Grund wurde bei der Datenanalyse mit besonderer Vorsicht vorgegangen. Falls es geringste Zweifel gab, wurde der Sachverhalt nicht als Fehler eingestuft. Nichtsdestotrotz konnten in der Stichprobe 28 Groß- und Kleinschreibfehler festgestellt werden. Somit handelt es sich bei den Groß- und Kleinschreibfehler mit einem Anteil von 25,5 % der Gesamtfehlerzahl um die zweithäufigste Fehlerart. Bei mehr als zwei Dritteln der Fehler handelt es sich um die falsche Großschreibung der Substantive, in den restlichen Fällen liegt eine falsche Kleinschreibung der Namen vor. Da die Groß- und Kleinschreibung bei der Indexierung nicht berücksichtigt wird, sind Fehler dieser Art nicht retrievalrelevant.

Die dritthäufigste Fehlerart sind Auslassungen mit insgesamt 17 Fällen und einem Anteil von 15,5 % von der Gesamtzahl der festgestellten Fehler. Fast zwei Drittel der Auslassungen entfallen dabei auf den russischen Buchstaben „б“. Die Auslassung dieses Buchstabens anstatt seiner Verwechslung könnte ein Indiz dafür sein, dass die entsprechenden Datensätze von Personen mit Russischkenntnissen aufgenommen wurden. Auslassung gehört zu den Fehlerarten, die das Retrieval negativ beeinflussen können.

Der Anteil der Weglassung der diakritischen Zeichen an der Gesamtfehlerzahl liegt mit insgesamt 15 festgestellten Fehlern bei 13,6 % und befindet sich somit im mittleren Bereich auf Platz vier. Am häufigsten werden diakritische Zeichen beim Buchstaben „ё“ ausgelassen, was eventuell auf bestimmte sprachliche Hintergründe zurückzuführen ist. Die Weglassung der diakritischen Zeichen ist in der Praxis nicht retrievalrelevant, da diakritische Zeichen bei Indexierung nicht berücksichtigt werden und nur der Grundbuchstabe indexiert wird.

Während der Datenanalyse wurden 8 Datensätze festgestellt, in denen ein anderer Transliterationsstandard als die vorgesehene Norm DIN 1460 verwendet wurde. Somit liegt der Anteil dieser Fehlerart an der Gesamtfehlerzahl bei 7,3 %. In allen Datensätzen wurden dabei fälschlicherweise Buchstabenkombinationen aus dem Transliterationsstandard der Library of Congress und der American Library Association verwendet. Dies könnte darauf hindeuten, dass während der Katalogisierung amerikanische Fremddaten als Grundlage hinzugezogen wurden. Generell sind die Fehler dieser Art retrievalrelevant.

In der Stichprobe kam es zu 6 Fällen des Einfügens. Der Anteil dieser Fehlerart an der Gesamtfehlerzahl beträgt somit 5,5 %. Da die Fehlerzahl und die Häufigkeit der einzelnen fälschlicherweise eingefügten Buchstaben niedrig sind, können in Bezug auf diese Fehlerkategorie keine Muster festgestellt werden. Generell lässt sich festhalten, dass diese Fehlerart einen negativen Einfluss auf das Retrieval hat.

Die Fehlerart „Verwendung falscher diakritischer Zeichen“ war in der Formulierung der möglichen Fehlerarten ursprünglich nicht vorgesehen. Während der Datenanalyse wurde jedoch entschieden, derartige Fehler als eine gesonderte Fehlerart zu behandeln. Insgesamt wurden in der Stichprobe drei Mal falsche diakritische Zeichen verwendet. In allen Fällen handelte es sich um ein Breve anstatt eines Hatscheks. Der Anteil dieser Fehlerart an der Gesamtfehlerzahl liegt bei 2,7 % und ist somit gering. Da in allen drei Fällen ein jeweils anderer Grundbuchstabe betroffen war und die Fehlerzahl niedrig ist, kann für diese Fehlerart kein Muster identifiziert werden. Diese Fehlerart ist ebenso wie das Weglassen diakritischer Zeichen nicht retrievalrelevant.

In der Stichprobe wurden lediglich zwei Doppelungen festgestellt. Ihr Anteil an der Gesamtzahl der Fehler liegt bei 1,8 % und ist somit eher unbedeutend. Wegen der geringen Zahl der festgestellten Doppelungen können keine Muster ihrer Entstehung identifiziert werden. Es wird vermutet, dass es sich um reine Flüchtigkeitsfehler handelt, die jedoch einen negativen Einfluss auf das Retrieval haben.

In der Stichprobe wurde nur eine einzige Vertauschung festgestellt. Somit liegt der Anteil dieser Fehlerart an der Gesamtzahl der festgestellten Fehler bei 0,9 % und ist minimal. Vermutlich handelt es sich dabei ebenfalls um einen Flüchtigkeitsfehler. Diese Fehlerart ist jedoch retrievalrelevant.

Die Zusammenfassung der Fehlerarten und ihrer Anteile an der Gesamtfehlerzahl nach ihrer Auswirkung auf das Retrieval ergibt, dass ca. 60 % der Transliterationsfehler potenziell einen negativen Einfluss auf das Retrieval haben und das Auffinden der jeweiligen Monografie durch die Nutzer*innen beeinträchtigen könnten. Die restlichen ca. 40 % der Transliterationsfehler sind hingegen nicht retrievalrelevant.

3.4 Mögliche Hilfestellungen

Schwierigkeiten, die bei der Katalogisierung fremdsprachiger Medien entstehen, sind bei Weitem nicht neu. Es wurde bereits eine Reihe von möglichen Hilfestellungen formuliert, um diese Aufgabe zu erleichtern und die Qualität der Erschließung zu verbessern.

In seinem Werk über Fehlinformationen in kooperativ gepflegten bibliothekarischen Katalogen geht Bade unter anderem ausführlich auf sog. linguistische Fehler ein. Darunter versteht er Fehler, die bei der Erschließung von Medien entstehen, welche in einer Sprache vorliegen, die der oder die Erschließende nicht kennt, und die dementsprechend auf die mangelnden Sprachkenntnisse zurückzuführen sind.⁹⁴

Zur Lösung dieses Problems schlägt Bade folgende Maßnahmen vor:⁹⁵

- neue Sprachen lernen. Die Erweiterung eigener Kenntnisse nicht nur auf neue Themengebiete, sondern auch auf neue Sprachen sollte Teil der kontinuierlichen

⁹⁴ Vgl. Bade, David (2002): The creation and persistence of misinformation in shared library catalogs, S. 6.

⁹⁵ Vgl. ebd., S. 11.

Weiterbildung der in Erschließung tätigen Bibliothekarinnen und Bibliothekare sein;

- warten, bis in der kooperativen Datenbank ein „guter“ Datensatz angelegt wird;
- mit der Erschließung externe Personen oder Organisationen beauftragen, die über entsprechende Sprachkenntnisse verfügen;
- mit anderen Bibliotheken kooperieren.

Zur Vermeidung von linguistischen Fehlern soll laut Bade folgendes unterlassen werden:⁹⁶

- neue Datensätze anlegen, wenn der oder die Erschließende sich nicht sicher ist, was er oder sie tut;
- regelmäßig Normdatensätze anlegen, ohne über die Kenntnisse der jeweiligen Sprache zu verfügen;
- ohne ausreichende Begründung Schlagwörter zum bereits vorhandenen Datensatz hinzufügen.

Es ist kaum anzuzweifeln, dass die Umsetzung von Bades Vorschlägen zu einer Minimierung der linguistischen Fehler in bibliothekarischen Katalogen führen würde. Die meisten Maßnahmen können aber in der Praxis schwer umgesetzt werden. Kaum eine Bibliothekarin oder ein Bibliothekar hat die Zeit und genug Motivation, eine neue Sprache zu lernen, nur weil sie oder er ab und zu ein fremdsprachiges Buch katalogisieren muss. Mit sinkenden Etats und immer wieder vorkommenden Einsparungen sowie dem wachsenden Fachkräftemangel hat kaum eine Bibliothek die Möglichkeit, Personal mit spezifischen Sprachkenntnissen einzustellen. In dieser Situation sind Erschließende gezwungen, selbstständig Datensätze für fremdsprachige Medien anzulegen, obwohl sie nicht viel von der Sprache verstehen. Das Hoffen darauf, dass irgendjemand mit besseren Sprachkenntnissen einen Datensatz anlegt, kann scheitern, da die Bestände in der jeweiligen Sprache im jeweiligen Verbund gering sein können. Diese Strategie hätte zur Folge, dass ein Buch oder ein anderes Medium einfach in einem Schrank liegen bleibt und langsam in Vergessenheit gerät, was für Bibliotheksnutzer*innen nicht hilfreich sein würde. Wie die Erfahrungen der Verfasserin während ihres Praktikums im Staatlichen Museum für Naturkunde Stuttgart zeigten, kommen solche Situationen in der Praxis durchaus vor.

Für Bibliothekarinnen und Bibliothekare, die sich zumindest grundlegende Kenntnisse der russischen Sprache aneignen möchten, existieren sogar spezielle Lernmaterialien. So möchte Rosalind Kent mit ihrer Russisch-Lernhilfe Bibliothekarinnen und Bibliothekare unterstützen, die entweder gar kein Russisch gelernt oder den Unterricht abgebrochen haben. Sie hält es für wichtig, zuerst das russische Alphabet auswendig zu lernen und bietet dazu ein paar hilfreiche Tipps, die das Lernen erleichtern. Danach folgt die

⁹⁶ Vgl. Bade, David (2002): The creation and persistence of misinformation in shared library catalogs, S. 11.

Erläuterung der wichtigsten Grammatikfragen. Den Lernenden wird vorgeschlagen, mithilfe eines Wörterbuches einzelne Titel- und Verantwortlichkeitsangaben sowie vollständige Titelblätter zu übersetzen.⁹⁷

Das kleine speziell für Bibliothekarinnen und Bibliothekare angepasste Lehrbuch mag für diejenigen eine gute Hilfe sein, die schon gewisse Vorkenntnisse der russischen Sprache haben und diese auffrischen möchten. Für die meisten Personen, die nie Russischunterricht hatten, wäre die Herausforderung, selbstständig einige Aspekte der Sprache zu erlernen, einfach zu groß.

Angeichts der aufgeführten Überlegungen wäre eine bessere Kooperation zwischen Bibliotheken sinnvoll, denn in einem Verbund lassen sich mit hoher Wahrscheinlichkeit Fachkräfte mit ausreichenden Kenntnissen der russischen Sprache finden. Ähnlich zu den Fachredaktionen im Bereich Schlagwörter wären Fachredaktionen in Bezug auf bestimmte Sprachen denkbar. Die mit der Fachredaktion beauftragten Personen könnten im Verbund als Kontaktpersonen bei Fragen auftreten sowie die neu angelegten Datensätze in den jeweiligen Sprachen kontrollieren. Das Ausüben solcher Funktionen würde allerdings einen größeren Zeitaufwand mit sich bringen und somit auch Kosten für die betroffenen Bibliotheken erhöhen. In diesem Bezug wäre eine zentrale Entscheidung erforderlich, ob die entstehenden Kosten angesichts des damit verbundenen Mehrwertes hingenommen und verteilt werden sollen.

Die Ergebnisse dieser Bachelor-Arbeit liefern ebenfalls Erkenntnisse, die zur Minderung der Transliterationsfehler beitragen können. Im Rahmen der Datenanalyse wurden einige Fehlermuster festgestellt und somit die häufigen Problemstellen in der Praxis identifiziert. Da fast zwei Drittel der Auslassungen auf den russischen Buchstaben „Ъ“ entfallen, sollte diesem Buchstaben besondere Aufmerksamkeit gewidmet werden. Aufpassen sollte man auch bei dem Buchstaben „ё“ (russisches „э“), bei dem diakritische Zeichen besonders oft weggelassen werden. Bei der Verwendung der diakritischen Zeichen sollte außerdem generell auf das Auswählen richtiger diakritischer Zeichen in Übereinstimmung mit DIN 1460 geachtet werden. Bei der Übernahme von Fremddaten ist besondere Vorsicht in Bezug auf die darin enthaltene Transliteration geboten, da in Fremddaten abweichende Transliterationsstandards verwendet werden können.

Die Datenanalyse zeigte, dass die Groß- und Kleinschreibung sowie die Erkennung des Titelnachsatzes insbesondere problematisch sind, wenn die katalogisierenden Personen keine Kenntnis der russischen Sprache haben. Hier könnte die im Punkt 3.2.4.6. erwähnte Annotationskarte auf der Rückseite des Titelblattes behilflich sein. Sie enthält bibliografische Informationen zum jeweiligen Werk, die im Gegensatz zum frei gestalteten Titelblatt in standardisierter Form aufgeführt sind.

⁹⁷ Vgl. Kent, Rosalind (1974): Reading the Russian language: a guide for librarians and other professionals. Marcel Dekker, Inc., New York

4 Fazit und Ausblick

In der vorliegenden Bachelor-Arbeit wurden Transliterationsfehler bei der Formalerschließung russischsprachiger Monografien in der Verbunddatenbank K10plus untersucht. Der Anteil der festgestellten Datensätze, die Transliterationsfehler enthielten, belief sich auf 14,5 % der Gesamtzahl der geprüften Datensätze und ist somit nicht übermäßig hoch. Am häufigsten kommen die Fehlerarten Ersetzung und falsche Groß- und Kleinschreibung vor, gefolgt von Auslassung auf Platz drei. Verwendung falscher diakritischer Zeichen, Vertauschung und Doppelung sind hingegen sehr selten. Die restlichen Fehlerarten Weglassung der diakritischen Zeichen, Verwendung eines falschen Transliterationsstandards sowie Einfügen befinden sich im mittleren Bereich.

Im Rahmen der Bachelor-Arbeit wurden Forschungsfragen formuliert, die anhand der durchgeführten Datenanalyse auch beantwortet werden konnten. In Bezug auf einige Fehlerarten war es nicht möglich, bestimmte Muster zu identifizieren und belastbare Schlussfolgerungen zu ziehen. Grund dafür waren meistens entweder der Charakter der festgestellten Fehler oder zu niedrige Fehlerzahl. In Bezug auf andere Fehlerarten könnten hingegen Muster und Problemstellen festgestellt werden. Diese Erkenntnisse können zur Minderung der Transliterationsfehler in der Praxis verwendet werden.

Die Unterteilung der Fehlerarten nach ihrer möglichen Auswirkung auf das Retrieval⁹⁸ ergibt, dass ca. 60 % der Transliterationsfehler potenziell einen negativen Einfluss auf das Retrieval haben. Diese Vermutung wurde im Rahmen der vorliegenden Bachelor-Arbeit nicht überprüft. Die vorgestellten Ergebnisse ließen sich demnach durch eine Untersuchung ergänzen, ob und wie sich die einzelnen Tippfehlerarten auf das Retrieval auswirken.

In der durchgeführten Analyse wurden bewusst Datensätze mit Originalschrift nicht berücksichtigt, da dem Erfassen der Daten in Originalschrift mit anschließender Transliteration mithilfe des WinIBW-Tools ein ganz anderer Ansatz als der manuellen Transliteration zugrunde liegt. Da dieser Umstand eine Verzerrung der Analyseergebnisse verursachen könnte, wurden Datensätze mit Originalschrift schon bei der Formulierung der Suchanfrage ausgeschlossen. In diesem Zusammenhang wäre es lohnenswert zu untersuchen, ob Datensätze mit Originalschrift weniger Tippfehler aufweisen als die Datensätze mit manueller Transliteration.

Anhand der Analyseergebnisse wurde eine Fehlerquote in Höhe von 14,5 % berechnet, die als mäßig beurteilt wurde. Für die weitere Forschung wäre interessant, die durchgeführte Analyse durch einen Vergleich der Tippfehleranfälligkeit zwischen transliterierten Datensätzen und Datensätzen ohne jegliche Form der Transliteration (zum Beispiel bei

⁹⁸ Vgl. Bibliotheksservice-Zentrum Baden-Württemberg (2019): Tipps und Hinweise zur Recherche.

deutschsprachigen Medien) zu ergänzen. Dazu wäre es notwendig, die Tippfehler anhand einer ähnlich großen Stichprobe der Datensätze mit dem entsprechenden Sprachkode für Deutsch zu analysieren.

Somit liefert die vorliegende Bachelor-Arbeit interessante Erkenntnisse über Transliterationsfehler bei der Formalerschließung russischsprachiger Monografien, die als Grundlage für die weitere Forschung zu diesem Thema verwendet werden können.

Literaturverzeichnis

- Agenbroad, James E. (2006): Romanization Is Not Enough, in: Cataloging & Classification Quarterly 42, H. 2, S. 21–34. URL: https://www.tandfonline.com/doi/abs/10.1300/J104v42n02_03 (letzter Zugriff: 26.02.2020)
- Albrecht, Rita; Wiesner, Hans-Jörg, Hrsg. (2011): Umschriften in Bibliotheken. Anwendung, Verfahren, Tabellen. Berlin, Beuth
- Auswärtiges Amt (2019): Belarus: Überblick. URL: <https://www.auswaertiges-amt.de/de/aussenpolitik/laender/belarus-node/belarus/201902?openAccordionId=item-201914-0-panel> (letzter Zugriff: 26.02.2020)
- Bade, David (2002): The creation and persistence of misinformation in shared library catalogs. Language and subject knowledge in a technological era. Champaign (Ill.), Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign
- Ballard, Terry (1992): Spelling and Typographical Errors in Library Databases, in: Computers in Libraries 12, H. 5, S. 14–19
- Ballard, Terry (2008): Systematic Identification of Typographical Errors in Library Catalogs, in: Cataloging & Classification Quarterly 46, H. 1, S. 27–33. URL: <https://doi.org/10.1080/01639370802182976> (letzter Zugriff: 26.02.2020)
- Beall, Jeffrey (1991): Ideas: The Dirty Database Test, in: American Libraries 22, H. 3, S. 197
- Beall, Jeffrey (2004): Using OCLC Connexion to find typographical errors in authority records, in: OCLC Systems & Services: International digital library perspectives 20, H. 2, S. 71–75. URL: <https://www.emerald.com/insight/content/doi/10.1108/10650750410539077/full/html> (letzter Zugriff: 26.02.2020)
- Beall, Jeffrey; Kafadar, Karen (2004): The Effectiveness of Copy Cataloging at Eliminating Typographical Errors in Shared Bibliographic Records, in: Library Resources & Technical Services 48, H. 2, S. 92–101. URL: <https://journals.ala.org/index.php/lrts/article/view/5417/6652> (letzter Zugriff: 26.02.2020)
- Beall, Jeffrey; Kafadar, Karen (2007): Measuring Typographical Errors' Impact on Retrieval in Bibliographic Databases, in: Cataloging & Classification Quarterly 44, H. 3-4, S. 197–211. URL: https://www.tandfonline.com/doi/abs/10.1300/J104v44n03_03 (letzter Zugriff: 26.02.2020)
- Bibliotheksservice-Zentrum Baden-Württemberg (2019): Tipps und Hinweise zur Recherche. URL: <https://wiki.bsz-bw.de/doku.php?id=v-team:faq:rechtipps> (letzter Zugriff: 26.02.2020)

- Bologov, Petr (2013): Vse, ja teper' Sis'kin!, URL: <https://lenta.ru/articles/2013/05/21/names/> (letzter Zugriff: 26.02.2020)
- Bourne, Charles P. (1977): Frequency and impact of spelling errors in bibliographic data bases, in: Information processing & management 13, H. 1, S. 1–12
- Brewer, Michael M. (2009): Romanization of Cyrillic Script: Core Competencies and Basic Research Strategies for Slavic Students, Scholars, and Educators, in: Slavic & East European Information Resources 10, H. 2-3, S. 244–256. URL: <https://www.tandfonline.com/doi/abs/10.1080/15228880903115128> (letzter Zugriff: 26.02.2020)
- Bußmann, Hadumod, Hrsg. (2008): Lexikon der Sprachwissenschaft. 4. Aufl. Kröner, Stuttgart
- Cherukodan, Surendran et al. (2013): Analytical Study of Typographical Errors in OPACS and Corrective Measures, in: Arora, Jagdish; Gupta, Pawan K. (Hrsg.): Library vision 2020. Moving towards the future: 9th Convention CALIBER - 2013, Convention on Automation of Libraries in Education and Research : March 21-23, 2013 at INFLIBNET Centre, Gandhinagar, Gujarat, India. Gandhinagar, Information and Library Network Centre, Gandhinagar, S. 116–123. URL: <http://ir.inflibnet.ac.in/bitstream/1944/1755/1/14.pdf> (letzter Zugriff: 26.02.2020)
- Cour européenne des Droits de l'Homme (2001): L'affaire Siskina and Siskins v. Latvia, requête no 59727/00, URL: <http://hudoc.echr.coe.int/eng?i=001-43126> (letzter Zugriff: 26.02.2020)
- Cour européenne des Droits de l'Homme (2004): L'affaire Kuharec alias Kuhareca c. Lettonie, requête no 71557/01, URL: <http://hudoc.echr.coe.int/eng?i=001-68203> (letzter Zugriff: 26.02.2020)
- Deutsche Nationalbibliothek (2007): Regeln für die alphabetische Katalogisierung in wissenschaftlichen Bibliotheken. Elektronische Ausgabe basierend auf der 2., überarbeiteten Auflage, Stand: April 2006. URL: https://files.dnb.de/pdf/rak_wb_netz.pdf (letzter Zugriff: 26.02.2020)
- DIN (1962): DIN 1460 Transliteration slawischer kyrillischer Buchstaben. Berlin, Köln, Beuth
- DIN (2019): Norm-Entwurf DIN 1460. Umschrift kyrillischer Alphabete - Umschrift kyrillischer Alphabete slawischer Sprachen. URL: <https://www.din.de/de/mitwirken/entwuerfe/ne-stellung/wdc-beuth:din21:313832710> (letzter Zugriff: 26.02.2020)
- Duden (o. J.): Umschrift. URL: <https://www.duden.de/rechtschreibung/Umschrift> (letzter Zugriff: 26.02.2020)
- Ermolovič, D.I. (2012) Šiškin protiv Siskina, URL: <http://yermolovich.ru/index/0-38> (letzter Zugriff: 26.02.2020)

- European Court of Human Rights (2004) Case of Mentzen v. Latvia, Application no. 71074/01, URL: <http://hudoc.echr.coe.int/eng?i=001-70407> (letzter Zugriff: 26.02.2020)
- European Court of Human Rights (2007): Case of Bulgakov v. Ukraine, Application no. 59894/00, URL: <http://hudoc.echr.coe.int/eng?i=001-82241> (letzter Zugriff: 26.02.2020)
- Federal'noe agentstvo po techničeskomu regulirovaniju i metrologii (2006): GOST R 7.0.4.-2006 Izdanija. Vychodnye svedenija. Obščie trebovanija i pravila oformlenija. Moskau, Standartinform
- Franceschini, Fiorenzo; Maisano, Domenico; Mastrogiacomo, Luca (2016): Empirical analysis and classification of database errors in Scopus and Web of Science, in: Journal of Informetrics 10, H. 4, Seite 933-953. URL: <https://doi.org/10.1016/j.joi.2016.07.003> (letzter Zugriff: 26.02.2020)
- Gantert, Klaus (2016): Bibliothekarisches Grundwissen. 9. Aufl. Berlin, Boston, De Gruyter Saur
- Gentner, Donald R. et al. (1983): A glossary of terms including a classification of typing errors, in: Cooper, William E. (Hrsg.): Cognitive aspects of skilled typewriting. Berlin, Springer, S. 39–44
- Glück, Helmut, Hrsg; Rödel, Michael, Hrsg. (2016): Metzler Lexikon Sprache. 5. Aufl. J.B. Metzler Verlag, Stuttgart
- Gramota.ru (o. J.): Pravila russkoj orfografii i punktuacii 1956. Orfografija. Propisnye bukvy. URL: <http://new.gramota.ru/biblio/readingroom/rules/139-prop> (letzter Zugriff: 26.02.2020)
- Heinz, Christof; Kuße, Holger, Hrsg. (2015): Slawischer Sprachvergleich für die Praxis. Frankfurt a.M., Peter Lang
- ISO (1995): Information and documentation - transliteration of Cyrillic characters into Latin characters - Slavic and non-Slavic languages. 2. ed. Genève
- Kalinin, S. Ju. (2003): Kak pravil'no oformit' vychodnye svedenija izdanija. Posobie dlja izdatelja. Moskau, Ékonomist"
- Kempgen, Sebastian, Hrsg. (2009): Die slavischen Sprachen. Ein internationales Handbuch zu ihrer Struktur, ihrer Geschichte und ihrer Erforschung. Berlin, New York, de Gruyter
- Kent, Rosalind (1974): Reading the Russian language: a guide for librarians and other professionals. Marcel Dekker, Inc., New York
- Klauß, Henning (2001): Tippfehler in Bibliothekskatalogen. Möglichkeiten einer EDV-gestützten Ermittlung, in: Bibliotheksdienst 35, H. 7/8, S. 868–876
- Library of Congress (2012): ALA-LC Romanization Tables. URL: <https://www.loc.gov/catdir/cpsol/romanization/russian.pdf> (letzter Zugriff: 26.02.2020)
- Lopatin, V. V., Hrsg. (2009): Pravila russkoj orfografii i punktuacii. Polnyj akademičeskij spravočnik. URL: <http://orthographia.ru/> (letzter Zugriff: 26.02.2020)

- Mangold, Max (1965): Transliteration und Transkription. Mannheim, Dudenverlag
- Matthews, Joseph R.; Hinchliffe, Lisa (2018): The evaluation and measurement of library services. Santa Barbara, California, Denver, Colorado, Libraries Unlimited TM
- Mežgosudarstvennyj sovet po standartizacii, metrologii i sertifikacii (2000): GOST 7.79-2000 (ISO 9-95). Pravila transliteracii kirillovskogo pis'ma latinskim alfavitom. Oficial'noe izdanie 2002. Minsk, IPK Izdatel'stvo standartov
- Mežgosudarstvennyj sovet po standartizacii, metrologii i sertifikacii (1998): GOST 7.51-98. Kartočki dlja katalogov i kartotek. Katalogizacija v izdanii. Sostav, struktura dannyh i isdatel'skoe oformlenie. Minsk
- o. A. (1966): Instruktionen für die alphabetischen Kataloge der preussischen Bibliotheken vom 10. Mai 1899. Zweite Ausgabe in der Fassung vom 10. August 1908, unveränderter Nachdruck. Wiesbaden, Harrassowitz
- o. A. (1975): Regeln für die alphabetische Katalogisierung. Anlage 5. Tabellen für die Umschrift nichtlateinischer Schriftzeichen in die Buchstaben der lateinischen Schrift. Vorabdruck. München
- o. A. (1977): Regeln für die alphabetische Katalogisierung. Wiesbaden, Reichert
- o. A. (2016): Typo of the day for libraries - Index. URL: <http://libtypos.pbworks.com/w/page/17113321/FrontPage> (letzter Zugriff: 26.02.2020)
- o. A. (2017): Typo of the day for libraries. URL: <http://librarytypos.blogspot.com/> (letzter Zugriff: 26.02.2020)
- Poll, Roswitha; Boekhorst, Peter te (2007): Measuring quality. Performance measurement in libraries. 2. Aufl. München, K.G. Saur
- Randall, Barbara Nichols (1999): Spelling Errors in the Database: Shadow or Substance?, in: Library Resources & Technical Services 43, H. 3, S. 161–169. URL: <https://doi.org/10.5860/lrts.43n3.161> (letzter Zugriff: 26.02.2020)
- RDA-Toolkit. URL: <http://access.rdatoolkit.org/> (letzter Zugriff: 26.02.2020)
- Rehder, Peter, Hrsg. (2009): Einführung in die slavischen Sprachen. Mit einer Einführung in die Balkanphilologie. 6. Aufl. Darmstadt, WBG
- Ryans, Cynthia C. (1978): A study of errors found in non-MARC cataloging in a machine-assisted system, in: Journal of library automation 11, H. 2, S. 125–132
- SWB-GBW (2019): Erfassung nicht-lateinischer Schriften. URL: https://opus.k10plus.de/frontdoor/deliver/index/docId/395/file/K10plus_Nichtlateinische_Schriften.pdf (letzter Zugriff: 26.02.2020)
- Tüür-Fröhlich, Terje (2018): Eine „autoritative“ Datenbank auf dem Prüfstand: Der Social Sci-ences Citation Index (SSCI) und seine Datenqualität, in: Information. Wissenschaft & Praxis 69, H. 5–6, S. 265–275. URL: <https://doi.org/10.1515/iwp-2018-0050> (letzter Zugriff: 26.02.2020)

Wellisch, Hans H. (1978): The conversion of scripts - its nature, history, and utilization.
New York, Wiley

Wiktionary (2018): Umschrift. Version vom 15.11.2018. URL:
<https://de.wiktionary.org/wiki/Umschrift> (letzter Zugriff: 26.02.2020)

Anhang A. Analyseprotokoll

Das Analyseprotokoll wird in der elektronischen Version der vorliegenden Bachelor-Arbeit bereitgestellt.